```
   1 CTGCAGCCAA CTTTGTTGAC CATCTCCGCA ATGCCTTGGA CGTCCTGCAT
  51 AGAGAGCTTT TCCCTTAGGT GCCCAGAGTC CTGGTCAACC TCGTGGACTT
 101 CCTGAACCCC ACTATCATGC GGCAGGTGTT CCTGGGAAAC CCAGACAAGT
 151 GCCCAGTGCA GCAGGCCAGA GCAGCATGCG CGAGCTGGTG GGGTCAGGCC
 201 GCTATGACAC GCAGGAGGAC TTCTCTGTGG TGCTGCAGCC CTTCTTCCAG
 251 AACATCCAGC TCCCTGTCCT GGCGCTTGAA CCACTTGGAA GCAAAACAGA
 301 GACCCTGGAC CTGAGAGCAG AGATGCCCAT CACCTGTCCC ACTCAGAATG
 351 AGCCCTTCCT GAGAACCCCT CGGAATAGTA ACTACACGTA CCCCATCAAG
 401 CCAGCCATTG AGAACTGGGG CAGTGACTTC CTGTGTACAG AGTGGAAGGC
 451 TTCCAATAGT GTTCCAACCT CTGTCCACCA GCTCCGACCA GCAGACATCA
 501 AAGTGGTGGC CGCCCTGGGT GACTCTCTGA CTACAGCAGT GGGAGCTCGA
 551 CCAAACAACT CCAGTGACCT ACCCACATCT TGGAGGGGAC TCTCTTGGAG
 601 CATTGGAGGG GATGGGAACT TGGAGACTCA CACCACACTG CCCAACATTC
 651 TGAAGAAGTT CAACCCTTAC CTCCTTGGCT TCTCTACCAG CACCTGGGAG
 701 GGGACAGCAG GACTAAATGT GGCAGCGGAA GGGGCCAGAG CTAGGGACAT
 751 GCCAGCCCAG GCCTGGGACC TGGTAGAGCG AATGAAAAAC AGCCCCGACA
 801 TCAACCTGGA GAAAGACTGG AAGCTGGTCA CACTCTTCAT TGGGGTCAAC
 851 GACTTGTGTC ATTACTGTGA GAATCCGGAG GCCCACTTGG CCACGGAATA
 901 TGTTCAGCAC ATCCAACAGG CCCTGGACAT CCTCTCTGAG GAGCTCCCAA
 951 GGGCTTTCGT CAACGTGGTG GAGGTCATGG AGCTGGCTAG CCTGTACCAG
1001 GGCCAAGGCG GGAAATGTGC CATGCTGGCA GCTCAGAACA ACTGCACTTG
1051 CCTCAGACAC TCGCAAAGCT CCCTGGAGAA GCAAGAACTG AAGAAAGTGA
1101 ACTGGAACCT CCAGCATGGC ATCTCCAGTT TCTCCTACTG GCACCAATAC
1151 ACACAGCGTG AGGACTTTGC GGTTGTGGTG CAGCCTTTCT TCCAAAACAC
1201 ACTCACCCCA CTGAACGAGA GAGGGGACAC TGACCTCACC TTCTTCTCCG
1251 AGGACTGTTT TCACTTCTCA GACCGCGGGC ATGCCGAGAT GGCCATCGCA
1301 CTCTGGAACA ACATGCTGGA ACCAGTGGGC CGCAAGACTA CCTCCAACAA
1351 CTTCACCCAC AGCCGAGCCA AACTCAAGTG CCCCTCTCCT GAGAGCCCTT
1401 ACCTCTACAC CCTGCGGAAC AGCCGATTGC TCCCAGACCA GGCTGAAGAA
1451 GCCCCCGAGG TGCTCTACTG GGCTGTCCCA GTGGCAGCGG GAGTCGGCCT
1501 TGTGGTGGGC ATCATCGGGA CAGTGGTCTG GAGGTGCAGG AGAGGTGGCC
1551 GGAGGGAAGA TCCTCCAATG AGCCTGCGCA CTGTGGCCCT CTAGGCCCGG
1601 GGGTGGGTCC TCACCCTAAA CTCCCTATAG CCACTCTCTT CACCGCCCTC
1651 TGCCCCAGCC ACTCCCGGCC ACCAGGACAT GCTTCAATGC CTGGTGCCAT
1701 AGGAAGCCCA GGGGACAGTC ACAACTTCTT GGGGCCTGGG CTTCTTCCAG
1751 GCCTATGCTC CTGGAATGGA TACATTTAAA TAAAGTCCAA AGCTATTTTA
1801 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA
```

FEATURES:

5'UTR:	1 - 175
Start Codon:	176
Stop Codon:	1592
3'UTR:	1595

Homologous proteins:

Top BLAST Hits

	Score	E
CRA|18000004885276 /altid=gi|464376 /def=sp|Q05017|PHLX_RABIT P...	750	0.0
CRA|18000005150386 /altid=gi|3172337 /def=gb|AAC40129.1| (AF045...	682	0.0
CRA|18000005121266 /altid=gi|2696236 /def=dbj|BAA23813.1| (D636...	665	0.0
CRA|18000005181876 /altid=gi|7498717 /def=pir||T20655 hypotheti...	228	9e-59
CRA|87000001028586 /altid=gi|7332170 /def=gb|AAF60857.1| (AC024...	210	3e-53
CRA|18000005040393 /altid=gi|7508802 /def=pir||T26083 hypotheti...	205	7e-52
CRA|89000000196200 /altid=gi|7293699 /def=gb|AAF49069.1| (AE003...	200	3e-50
CRA|89000000199135 /altid=gi|7297015 /def=gb|AAF52285.1| (AE003...	189	6e-47
CRA|18000004979533 /altid=gi|7499049 /def=pir||T16060 hypotheti...	161	2e-38
CRA|18000005184633 /altid=gi|7506410 /def=pir||T24016 hypotheti...	152	9e-36
CRA|18000005184632 /altid=gi|7506411 /def=pir||T24015 hypotheti...	122	1e-26
CRA|18000005182912 /altid=gi|7500588 /def=pir||T21835 hypotheti...	119	6e-26
CRA|87000001028649 /altid=gi|7332235 /def=gb|AAF60922.1| (AC006...	111	2e-23

BLAST dbEST hits:

gi|2079883 /dataset=dbest /taxon=9606 ...	724	0.0
gi|11593761 /dataset=dbest /taxon=960...	670	0.0
gi|7037501 /dataset=dbest /taxon=9606...	654	0.0
gi|12241943 /dataset=dbest /taxon=96...	632	e-179
gi|10367787 /dataset=dbest /taxon=960...	575	e-161
gi|9969781 /dataset=dbest /taxon=960...	547	e-153
gi|7667765 /dataset=dbest /taxon=9606...	531	e-148
gi|12241345 /dataset=dbest /taxon=96...	519	e-145
gi|2080047 /dataset=dbest /taxon=9606 ...	468	e-129
gi|7946640 /dataset=dbest /taxon=960...	323	7e-86

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

Expression information from BLAST dbEST hits:

gi\|2079883 Mixed (melaonocyte, fetal heart, pregnant uterus)
gi\|11593761 Kidney
gi\|7037501 Whole blood
gi\|12241943 Lung, normal
gi\|10367787 brain glioblastoma
gi\|9969781 Prostate
gi\|7667765 Colon
gi\|12241345 Lung, normal
gi\|2080047 Mixed (melaonocyte, fetal heart, pregnant uterus)
gi\|7946640 Colon

Expression information from PCR-based tissue screening panels:
Leukocyte

FIGURE 1, page 2 of 2

```
  1 MRELVGSGRY DTQEDFSVVL QPFFQNIQLP VLALEPLGSK TETLDLRAEM
 51 PITCPTQNEP FLRTPRNSNY TYPIKPAIEN WGSDFLCTEW KASNSVPTSV
101 HQLRPADIKV VAALGDSLTT AVGARPNNSS DLPTSWRGLS WSIGGDGNLE
151 THTTLPNILK KFNPYLLGFS TSTWEGTAGL NVAAEGARAR DMPAQAWDLV
201 ERMKNSPDIN LEKDWKLVTL FIGVNDLCHY CENPEAHLAT EYVQHIQQAL
251 DILSEELPRA FVNVVEVMEL ASLYQGQGGK CAMLAAQNNC TCLRHSQSSL
301 EKQELKKVNW NLQHGISSFS YWHQYTQRED FAVVVQPFFQ NTLTPLNERG
351 DTDLTFFSED CFHFSDRGHA EMAIALWNNM LEPVGRKTTS NNFTHSRAKL
401 KCPSPESPYL YTLRNSRLLP DQAEEAPEVL YWAVPVAAGV GLVVGIIGTV
451 VWRCRRGGRR EDPPMSLRTV AL
```

FEATURES:
Functional domains and key regions:
[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 5

1	69-72	NYTY
2	127-130	NNSS
3	128-131	NSSD
4	289-292	NCTC
5	392-395	NFTH

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

386-389 RKTT

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 7

1	7-9	SGR
2	64-66	TPR
3	135-137	SWR
4	326-328	TQR
5	365-367	SDR
6	412-414	TLR
7	466-468	SLR

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 6

1	12-15	TQED
2	39-42	SKTE
3	56-59	TQNE
4	172-175	STWE
5	298-301	SSLE
6	326-329	TQRE

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 8
        1   123-128 GARPNN
        2   138-143 GLSWSI
        3   144-149 GGDGNL
        4   147-152 GNLETH
        5   179-184 GLNVAA
        6   278-283 GGKCAM
        7   441-446 GLVVGI
        8   445-450 GIIGTV

[6] PDOC00009 PS00009 AMIDATION
Amidation site

Number of matches: 2
        1   384-387 VGRK
        2   457-460 GGRR

[7] PDOC00016 PS00016 RGD
Cell attachment sequence

        349-351 RGD

[8] PDOC00200 PS00228 TUBULIN_B_AUTOREG
Tubulin-beta mRNA autoregulation signal

        1-4 MREL

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	164	184	0.694	Putative
2	432	452	1.956	Certain

FIGURE 2, page 2 of 7

>CRA|18000004885276 /altid=gi|464376 /def=sp|Q05017|PHLX_RABIT
          PHOSPHOLIPASE ADRAB-B PRECURSOR /dataset=nraa
          /length=1458
        Length = 1458

 Score = 750 bits (1915), Expect = 0.0
 Identities = 368/502 (73%), Positives = 407/502 (80%), Gaps = 33/502 (6%)

Query: 1   MRELVGSGRYDTQEDFSVVLQPFFQNIQLPVLA-------------------------- 33
             +RELV SGRYDT+EDFSVVLQPFF +IQLPVL
Sbjct: 955   LRELVESGRYDTREDFSVVLQPFFHSIQLPVLQDGRLDTSFFAPDCVHPNQKFHSQLSRA 1014

Query: 34   -----LEPLGSKTETLDLRAEMPITCPTQNEPFLRTPRNSNYTYPIKPAIENWGSDFLCT 88
                  LEPLG KT+ LDL A + +TCPTQNEPFLRT RNS+YTYP +PA+ENWGSDFLCT
Sbjct: 1015 LWRNMLEPLGGKTDALDLTAAITLTCPTQNEPFLRTFRNSDYTYPSRPAVENWGSDFLCT 1074

Query: 89   EWKASNSVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPTSWRGLSWSIGGDGN 148
             W AS VP SVH+L+P DIKVVAALGDSLT A+GARP+NSSD P WRGLSWSIGGDG
Sbjct: 1075 AWNASRGVPNSVHELQPGDIKVVAALGDSLTLMGARPSNSSDPPMFWRGLSWSIGGDGA 1134

Query: 149   LETHTTLPNILKKFNPYLLGFSTSTWEGTAGLNVAAEGARARDMPAQAWDLVERMKNSPD 208
             LETHTTLPNILKKFNP +LGFST T EGT GLNVA +GARA+DMPAQA DLVERM+NSP+
Sbjct: 1135 LETHTTLPNILKKFNPSILGFSTGTLEGTMGLNVAVQGARAQDMPAQARDLVERMRNSPE 1194

Query: 209   INLEKDWKLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVVEVM 268
             I+LEKDWKLVTLF+G NDLCH+CENPE   EYVQHIQQALD+L EELPR FVNVVEVM
Sbjct: 1195 IDLEKDWKLVTLFVGGNDLCHFCENPEGSSEGEYVQHIQQALDVLYEELPRTFVNVVEVM 1254

Query: 269   ELASLYQGQGGKCA-MLAAQNNCTCLRHSQSSLEKQELKKVNWNLQHGISSFSYWHQYTQ 327
             ELA L+Q QGG+CA +LAAQ++CTC ++SQSS+E QELKKVNWNLQ G+S SY HQY Q
Sbjct: 1255 ELAGLHQDQGGRCATLLAAQSHCTCFKYSQSSVEMQELKKVNWNLQSGLSRLSYSHQYVQ 1314

Query: 328   REDFAVVVQPFFQNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEMAIALWNNMLEPVGRK 387
             REDFAVVVQPFFQNTL PLN RGDTDLTFFS+DCFHFS+RGHAEMAIALWNNMLEPVG K
Sbjct: 1315 REDFAVVVQPFFQNTLVPLNGRGDTDLTFFSDDCFHFSERGHAEMAIALWNNMLEPVGHK 1374

Query: 388   TTSNNFTHSRAKLKCPSPESPYLYTLRNSRLLPDQAEEAPEVLYWAVPVAAGVGLVVGII 447
             TTSNNFT+SR KLKCPSP+SPYLYTLRNSRLLPDQAE P VLYWAVPVAAG GL++GI+
Sbjct: 1375 TTSNNFTYSRTKLKCPSPDSPYLYTLRNSRLLPDQAEADPTVLYWAVPVAAGAGLLIGIL 1434

Query: 448   GTVVWRCRRGGRREDPPMSLRT 469
              V R R   REDPP+SL T
Sbjct: 1435 AMVAGRGMRCRPREDPPLSLST 1456

 Score = 334 bits (847), Expect = 2e-90
 Identities = 191/445 (42%), Positives = 251/445 (55%), Gaps = 33/445 (7%)

Query: 4   LVGSGRYDTQEDFSVVLQPFFQNIQLPVLA-------------------LEPLGSKTETL 44
             L+ S +Y+TQE F+VV QPFF   L L   +EP+G K E
Sbjct: 264   LLASSKYNTQESFAVVFQPFFYESSLSALLAEPPLQDPTTLALSLWNRMMEPIGRKEEPF 323

Query: 45   DLRAEMPITCPTQNEPFLRTPRNSNY----TYPIKPAIENWGSDFLCTEWKASNSVPTSV 100
             +   P+ CPTQ P+L T RNS   + P   G++ C +   S+SVPTSV
Sbjct: 324   SEKERKPLRCPTQESPYLFTYRNSGQLTRVSQPQGKLEVREGTEIRCPDKDPSDSVPTSV 383

Query: 101   HQLRPADIKVVAALGDSLTTAVGA--RPNNSSDLPTSWRGLSWSIGGDGNLETHTTLPNI 158
             H+L+PADIKV+ A+GDSLT   GA +P N D+ T +RGLSWS+GGD N+ T TTL NI
Sbjct: 384   HRLKPADIKVIGAMGDSLTAGNGAGSQPGNILDVLTQYRGLSWSVGGDQNISTVTTLANI 443

Query: 159   LKKFNPYLLGFSTSTWEGT---AGLNVAAEGARARDMPAQAWDLVERMKNSPDINLEKDW 215
             L++FNP L GFS T   T   A N A GARA + QA LV MKN   IN ++DW

```
Sbjct:  444 LREFNPSLQGFSVGTGRETTSQAFFNQAVAGARADGLIPQAQRLVALMKNDTRINFQEDW 503

Query:  216 KLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVVEVME---LAS 272
            K++T+FIG NDLC +C +P +    + +I  ALDIL  E+PRAFVN+V+V+E     L
Sbjct:  504 KIITVFIGGNDLCDFCNDPVRYSPQNFTDNIGTALDILHAEIPRAFVNLVKVLEISKLRE 563

Query:  273 LYQGQGGKCAMLAAQNNCTC-LRHSQSSLEKQELKKVNWNLQHGISSFSYWHQYTQREDF 331
            LYQ    C + ++ C C L+  +S E   L +     Q        +Y  R+DF
Sbjct:  564 LYQETKVSCPRMILRSLCPCVLKFDDNSTEIASLIETIKEYQERTQQLIDSGRYDTRDDF 623

Query:  332 AVVVQPFFQNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEMAIALWNNMLEPVGRKTTSN 391
            VV+QPFF+    P + G  D +FF+ DCFHFS + HA  A ALWNNMLEPVG+KTT N
Sbjct:  624 TVVLQPFFEKVNMPKTQDGLPDNSFFAPDCFHFSSKAHAHAASALWNNMLEPVGQKTTHN 683

Query:  392 NFTHSRAKLKCPSPESPYLYTLRNS 416
            +F        + CP+   P+L T +NS
Sbjct:  684 DF-EGAVNITCPNQVWPFLSTYKNS 707


  Score =  323 bits (819),  Expect = 3e-87
  Identities = 181/456 (39%),  Positives = 261/456 (56%),  Gaps = 51/456 (11%)

Query:    2 RELVGSGRYDTQEDFSVVLQPFFQNIQLPVLA--------------------------- 33
            ++L+ SGRYDT++DF+VVLQPFF+ + +P
Sbjct:  609 QQLIDSGRYDTRDDFTVVLQPFFEKVNMPKTQDGLPDNSFFAPDCFHFSSKAHAHAASAL 668

Query:   34 ----LEPLGSKTETLDLRAEMPITCPTQNEPFLRTPRNSNYTYPIKPAIENWGSDFLCTE 89
                LEP+G KT  D   + ITCP Q  PFL T +NS          ++ +G+   C +
Sbjct:  669 WNNMLEPVGQKTTHNDFEGAVNITCPNQVWPFLSTYKNS---------VQGFGTWLPCRD 719

Query:   90 WKASNSVPTSVHQLRPADIKVVAALGDSLTTAVG--ARPNNSSDLPTSWRGLSWSIGGDG 147
             S S PTSVH LRPADI+VVAALGDSLT  +G  ++PN+ SD T +RGLS+S GGDG
Sbjct:  720 RSPSASPPTSVHALRPADIQVVAALGDSLTAGIGIGSKPNDLSDGTTQYRGLSYSSGGDG 779

Query:  148 NLETHTTLPNILKKFNPYLLGFSTSTWEGT---AGLNVAAEGARARDMPAQAWDLVERMK 204
            +L+  TTLPNIL++FN L+GF+ T + +   A  N A GA+ARD+ +Q   LV+RMK
Sbjct:  780 SLDNVTTLPNILRQFNSNLMGFAVGTGDASGTNAFFNQAVPGAKARDLMSQVQTLVQRMK 839

Query:  205 NSPDINLEKDWKLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNV 264
            +   +N ++DWK++T+ IG +DLC YC +   + A + H++ ALD L  E+PRA VN+
Sbjct:  840 DDHRVNFQEDWKVITVQIGASDLCDYCTDSNLYSAANFYDHLRDALDALHREVPRALVNL 899

Query:  265 VEVME---LASLYQGQGGKCAMLAAQNNCTC-LRHSQSSLEKQELKKVNWNLQHGISSFS 320
            V+ M    ++ G  KC + A   C C L  ++S E  L+ +    Q  +
Sbjct:  900 VDFMNPSVTRQVFLGNPDKCPVQQASALCNCVLSPRENSYELARLEALAQAYQSSLRELV 959

Query:  321 YWHQYTQREDFAVVVQPFFQNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEMAIALWNNM 380
            +Y  REDF+VV+QPFF +   P+ + G  D +FF+ DC H + + H++++ ALW NM
Sbjct:  960 ESGRYDTREDFSVVLQPFFHSIQLPVLQDGRLDTSFFAPDCVHPNQKFHSQLSRALWRNM 1019

Query:  381 LEPVGRKTTSNNFTHSRAKLKCPSPESPYLYTLRNS 416
            LEP+G KT + + T +   L CP+   P+L T RNS
Sbjct: 1020 LEPLGGKTDALDLT-AAITLTCPTQNEPFLRTFRNS 1054


  Score =  137 bits (341),  Expect = 3e-31
  Identities = 107/338 (31%),  Positives = 161/338 (46%),  Gaps = 42/338 (12%)

Query:   85 FLCTEWKASNSVPT-SVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPTSWRGLSWSI 143
            F C    + SVP+ SVH LRP+DIK VAA+G+ T      +     T R     +
Sbjct:   46 FPCDPKTLAESVPSESVHSLRPSDIKFVAAIGNVETAPDSGADDLEEQDGTEKRPEQACM 105

Query:  144 GGDGNLETHTTLPNILKKFNPY-LLGFSTSTWEGTAGLNVAAEGARARDMPAQAWDLVER 202
```

```
                G       T L +I+ +F+P  L+       T          +  G  A D+  QA +LV
Sbjct: 106  G------VVTVLSDIIGRFSPSALMPLCPET-------RLVPRGG-AEDLWMQATELVRS 151

Query: 203  MKNSPDINLEKDWKLVTLFIGVNDLCHYCENPEAH-LATEYVQHIQQALDILSEELPRAF 261
            M+ +P ++ E DWKL+ +F      C  C + +   L  + + + LD L +E+P+AF
Sbjct: 152  MRENPQLDFEHDWKLINVFFSNTSQCFPCPSAQQKGLVLGGMDKLTRTLDYLQQEVPKAF 211

Query: 262  VNVVEVMELASLYQGQGGKCAMLAAQNNCTCLRHSQSSLEKQELKKV--NWNLQHGISSF 319
            VN+V++ ELA+ + + G   + A  C CLR      E +L KV  W+      S
Sbjct: 212  VNLVDLSELAAFSRWRQG-AQLSPAAEPCRCLR------ETSQLTKVLTQWSYLEAWDSL 264

Query: 320  SYWHQYTQREDFAVVVQPFF-QNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEMAIALWN 378
              +Y  +E FAVV QPFF +++L+ L            +      +A++LWN
Sbjct: 265  LASSKYNTQESFAVVFQPFFYESSLSALLAEPPL-------------QDPTTLALSLWN 310

Query: 379  NMLEPVGRKTTSNNFTHSRAKLKCPSPESPYLYTLRNS 416
            M+EP+GRK    +   R  L+CP+ ESPYL+T RNS
Sbjct: 311  RMMEPIGRKEEPFS-EKERKPLRCPTQESPYLFTYRNS 347


>CRA|18000005150386 /altid=gi|3172337 /def=gb|AAC40129.1| (AF045454)
        phospholipase B [Cavia porcellus] /org=Cavia porcellus
        /taxon=10141 /dataset=nraa /length=1463
        Length = 1463

 Score =  682 bits (1741), Expect = 0.0
 Identities = 348/505 (68%), Positives = 389/505 (76%), Gaps = 38/505 (7%)

Query: 1    MRELVGSGRYDTQEDFSVVLQPFFQNIQLPVLA-------------------------- 33
            MRELV SGRYDT+EDFSVVLQPFF NI+LP+L
Sbjct: 954  MRELVESGRYDTREDFSVVLQPFFLNIRLPILEDGRPDTSFFAPDCINPGQKFHSQLSRA 1013

Query: 34   -----LEPLGSKTETLDLRAEMPITCPTQNEPFLRTPRNSNYTYPIKPAIENWGSDFLCT 88
                 LEP+GSKT+TLDL A++ + CPTQ EPFLRTP+NS+YTYP KPAIENWGSDFLCT
Sbjct: 1014 LWVNMLEPVGSKTDTLDLTADISLPCPTQEEPFLRTPQNSDYTYPTKPAIENWGSDFLCT 1073

Query: 89   EWKASNSVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPTSWRGLSWSIGGDGN 148
            EWK SNSVPTSVH+L+PADIKVVAALGDSLTTAVGAR +NSSDL  SWRGLSWSIGGDG
Sbjct: 1074 EWKPSNSVPTSVHKLQPADIKVVAALGDSLTTAVGARASNSSDLLMSWRGLSWSIGGDGA 1133

Query: 149  LETHTTLPNILKKFNPYLLGFSTSTWEGTAGLNVAAEGARARDMPAQAWDLVERMKNSPD 208
            LETHTTLPNILKKFNP + GFST T E TAG NVA E ARARDMPAQA DLVERMK S +
Sbjct: 1134 LETHTTLPNILKKFNPSIFGFSTGTLEETAGFNVAVEEARARDMPAQARDLVERMKASTE 1193

Query: 209  INLEKDWKLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVVE-V 267
            INLE DWKL+TLFIG NDLCHYC+NPE H A EYVQHI+QALDIL EELPRAF+NVV+ +
Sbjct: 1194 INLEMDWKLITLFIGSNDLCHYCDNPENHSAEEYVQHIRQALDILYEELPRAFINVVDII 1253

Query: 268  MELASLYQGQGGKC-AMLAAQNNCTCLRHSQSSLEKQELKKVNWNLQHGISSFSYWHQYT 326
            MELA L+QGQGG C A+L AQ+ C+CLRH  SS   QELKKV WNLQ +S  SY +YT
Sbjct: 1254 MELAGLHQGQGGHCTALLPAQSTCSCLRHFPSSVIQELKKVTWNLQSDMSRLSYQEKYT 1313

Query: 327  QREDFAVVVQPFFQNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEMAIALWNNMLEPVGR 386
            QREDFAVVVQPFFQNTL PL++ G TD TFFSEDC HFS+RGHAEMAIALWNNMLEPVG
Sbjct: 1314 QREDFAVVVQPFFQNTLIPLDKLGSTDPTFFSEDCLHFSERGHAEMAIALWNNMLEPVGH 1373

Query: 387  KTTSNNFTHSRAKLKCPSPESPYLYTLRNSRLLPDQAEEAPEVLYWAVPVAAG----VGL 442
            KTT NNFT++R KLKCPS ESPYLYTL+NS LP Q E+A V    V AA     VGL
Sbjct: 1374 KTTFNNFTYNRTKLKCPSTESPYLYTLQNSLSLPVQTEKASGVAPGIVSAAAAGGLLVGL 1433

Query: 443  VVGIIGTVVWRCRRGGRREDPPMSL 467
            +VGI+    +W  R  +++ PP S+
Sbjct: 1434 IVGILAVSLWSSFRRRQKKSPPESV 1458
```

```
Score =  348 bits (884), Expect = 7e-95
Identities = 199/442 (45%), Positives = 257/442 (58%), Gaps = 31/442 (7%)

Query: 4    LVGSGRYDTQEDFSVVLQPFFQNIQLPV----------LAL-------EPLGSKTETLDL 46
            L+ S ++ QE F+VV QPFF + PV          LAL       +P+G K E
Sbjct: 265  LLASSSFNDQESFAVVFQPFFYEVSSPVEEPPSQDPTTLALSLWNNMMKPVGQKDEPFST 324

Query: 47   RAEMPITCPTQNEPFLRTPRNSNYTYPI----KPAIENWGSDFLCTEWKASNSVPTSVHQ 102
            P+ CP+Q P+L T RNSNY + + E G++ C + S+S PTSVH+
Sbjct: 325  IERRPMKCPSQESPYLFTYRNSNYQSRLLKRQRQHKEREGTEIRCPDKDPSDSTPTSVHR 384

Query: 103  LRPADIKVVAALGDSLTTAVGA--RPNNSSDLPTSWRGLSWSIGGDGNLETHTTLPNILK 160
            L+PADIKV+ ALGDSLT  GA RP N D+ T +RGLSWSIG D N+ + TTLPNIL+
Sbjct: 385  LKPADIKVIGALGDSLTAGNGAGSRPGNILDVLTEYRGLSWSIGADHNISSVTTLPNILR 444

Query: 161  KFNPYLLGFSTSTWEGT---AGLNVAAEGARARDMPAQAWDLVERMKNSPDINLEKDWKL 217
            +FNP L GFST T +    A N A GARA D+ QA LV+ MKN   IN E+DWK+
Sbjct: 445  EFNPSLKGFSTGTGKANSVGAFFNQAVAGARAGDLIPQARTLVDLMKNHTSINFEEDWKI 504

Query: 218  VTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVVEVME---LASLY 274
            +T+FIG NDLC +C +P +    + +I+QALDIL  E+PRAFVN+V+V++   L LY
Sbjct: 505  ITVFIGGNDLCDFCSDPVTNSPENFTDNIRQALDILHAEVPRAFVNMVKVLQIVNLRELY 564

Query: 275  QGQGGKCAMLAAQNNCTC-LRHSQSSLEKQELKKVNWNLQHGISSFSYWHQYTQREDFAV 333
            +    C L +N C C L   +SE + L +N  Q        +Y  REDF V
Sbjct: 565  KDSRVSCPRLILRNLCRCVLLPDDNSTELESLIDINKKYQERTHQLIESGRYDTREDFTV 624

Query: 334  VVQPFFQNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEMAIALWNNMLEPVGRKTTSNNF 393
            V+QPFF+    P  G D T F+ DCFHFS + HA  A ALW NMLEPVG+KTT NNF
Sbjct: 625  VLQPFFEKVDIPKTSEGLPDNTSFAPDCFHFSSKTHARAASALWKNMLEPVGQKTTQNNF 684

Query: 394  THSRAKLKCPSPESPYLYTLRN 415
            +S  + CP+  PYL T +N
Sbjct: 685  ENS-IDIICPNQAFPYLSTYKN 705


Score =  314 bits (795), Expect = 2e-84
Identities = 178/455 (39%), Positives = 259/455 (56%), Gaps = 51/455 (11%)

Query: 3    ELVGSGRYDTQEDFSVVLQPFFQNIQLPVLA----------------------------- 33
            +L+ SGRYDT+EDF+VVLQPFF+ + +P +
Sbjct: 609  QLIESGRYDTREDFTVVLQPFFEKVDIPKTSEGLPDNTSFAPDCFHFSSKTHARAASALW 668

Query: 34   ---LEPLGSKTETLDLRAEMPITCPTQNEPFLRTPRNSNYTYPIKPAIENWGSDFLCTEW 90
               LEP+G KT   +  + I CP Q P+L T +N          IE G+  C E
Sbjct: 669  KNMLEPVGQKTTQNNFENSIDIICPNQAFPYLSTYKNG--------IEGHGTWLTCRER 719

Query: 91   KASNSVPTSVHQLRPADIKVVAALGDSLT--TAVGARPNNSSDLPTSWRGLSWSIGGDGN 148
            S S PTSVH LRPAD+++VVAALGDSLT  + +G++P + +D+ T +RGLS+S GGDG+
Sbjct: 720  TPSASPPTSVHALRPADVRVVAALGDSLTAGSGIGSKPGDLADVITQYRGLSYSSGGDGS 779

Query: 149  LETHTTLPNILKKFNPYLLGFSTSTWEGT---AGLNVAAEGARARDMPAQAWDLVERMKN 205
            L   TTLPNIL++FN  L G++ T + +    A LN A  GA+A ++ +Q  LV++MK+
Sbjct: 780  LMNVTTLPNILREFNSNLTGYAVGTGDASNTAFLNQAVPGAKAEELMSQVKTLVQKMKD 839

Query: 206  SPDINLEKDWKLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVV 265
            P IN +DWK++T+ IG NDLC++C + + + + +   H+  ALDIL  E+PRA VN+V
```

```
Sbjct: 840   DPRINFHEDWKVITVLIGTNDLCNHCTDLDLYSSANFFNHLLNALDILHREVPRALVNLV 899

Query: 266   EVME---LASLYQGQGGKCAMLAAQNNCTC-LRHSQSSLEKQELKKVNWNLQHGISSFSY 321
             + M   + ++G  KC + A    C C L  ++S E    +    Q  +
Sbjct: 900   DFMNPSIMRQVFLGNPDKCPVQQASILCNCVLSLRENSYELARMDALTRAYQSSMRELVE 959

Query: 322   WHQYTQREDFAVVVQPFFQNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEMAIALWNNML 381
             +Y  REDF+VV+QPFF N   P+ E G   D +FF+ DC +   + H++++ ALW NML
Sbjct: 960   SGRYDTREDFSVVLQPFFLNIRLPILEDGRPDTSFFAPDCINPGQKFHSQLSRALWVNML 1019

Query: 382   EPVGRKTTSNNFTHSRAKLKCPSPESPYLYTLRNS 416
             EPVG KT + + T +   L CP+ E P+L T +NS
Sbjct: 1020  EPVGSKTDTLDLT-ADISLPCPTQEEPFLRTPQNS 1053


  Score =  155 bits (389), Expect = 8e-37
  Identities = 110/351 (31%), Positives = 166/351 (46%), Gaps = 48/351 (13%)

Query: 85    FLCTEWKASNSVPT-SVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPTS---WRGLS 140
             F C+  K    ++P+ SVH L PADIK++AA+GD  T      N +   T   WRG
Sbjct: 45    FSCSPKKLGLNMPSESVHTLTPADIKLIAAIGDMETPPDSGAVNLDTSERTEKEPWRGCM 104

Query: 141   WSIGGDGNLETHTTLPNILKKFNPYLLGFSTSTWEGTAGLNVAAEGARARDMPAQAWDLV 200
             +         T L +I+ FNP +L +   W     AA    ++ QA +LV
Sbjct: 105   GMM---------TVLSDIISHFNPSVLLPTCPPWRS------AAVRGGVEELRTQAEELV 149

Query: 201   ERMKNSPDINLEKDWKLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQ---ALDILSEEL 257
             +K +P ++ ++DWKL+ +F    LC+ C  P AH     + ++ +     L  L +E+
Sbjct: 150   SSLKKNPQLDFQQDWKLINVFFSNASLCYLC--PSAHENGPLMSNMDKLAGILHYLHQEV 207

Query: 258   PRAFVNVVEVMELASLYQGQGGKCAMLAAQNNCTCLRHSQSSLEKQELKKVNWNLQHGIS 317
             PRAFVN+V++ E+ ++ +   G     +   C C       K +   + W+ Q
Sbjct: 208   PRAFVNLVDLFEVVAMPRWHQGTMLSRPSPEACGC----SGETSKLDTVVMQWSYQETWD 263

Query: 318   SFSYWHQYTQREDFAVVVQPFFQNTLTPLNERGDTDLTFFSEDCFHFSDRGHAEMAIALW 377
             S     +E FAVV QPFF    +P+ E    D T            +A++LW
Sbjct: 264   SLLASSSFNDQESFAVVFQPFFYEVSSPVEEPPSQDPT---------------TLALSLW 308

Query: 378   NNMLEPVGRKTTSNNFTHSRAKLKCPSPESPYLYTLRN----SRLLPDQAE 424
             NNM++PVG+K   + T R  +KCPS ESPYL+T RN    SRLL  Q +
Sbjct: 309   NNMMKPVGQKDEPFS-TIERRPMKCPSQESPYLFTYRNSYQSRLLKRQRQ 358
```

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00657	Lipase/Acylhydrolase with GDSL-like motif	158.4	1.6e-45	1
PF01347	Lipoprotein amino terminal region	1.9	6.1	1
CE00543	CE00543 steroid_receptor_N10	-0.8	1.5	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
PF01347	1/1	155	172	..	636	653	..	1.9	6.1
PF00657	1/1	110	233	..	1	146	[]	158.4	1.6e-45
CE00543	1/1	230	254	..	456	480	..	-0.8	1.5

FIGURE 2, page 7 of 7

```
   1 ATTCTGCAGC CAACTTTGTT GACCATCTCC GCAATGCCTT GGACGTCCTG
  51 CATAGAGAGG TGGGTGGGGG GCTTCCACAA GCTGGTAACA GCTCAAGCAT
 101 GGTGAGGGTG AAGGTGGATG GGGGGAAAGA ATGAGAGAAG AACCCCTTTC
 151 TCTCAAGGAG ACAGCCAAGG GCATGGANNN NNNNNNNNNN NNNNNNNNNN
 201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NGTGCCATTG
 951 CCGCTGCAGC CCCTTTGGGT GGCACCCATG GAGTTGTGTG ATGTACGGCC
1001 TGTAAGGCCT TACGAGGTAG CCCTGTATAG ACTCCTCCCC AGAACTCAAC
1051 TCCAGAAAGA CCAAGCTGGA TTGCTAAAGG AACCCATTCC TAGGGGCCCT
1101 GAGACAGCCC CAGGAAGAAG TGCCTGGAGC CCCCCTCTCA TCTGCAGCTT
1151 TTCCCTTAGG TGCCCAGAGT CCTGGTCAAC CTCGTGGACT TCCTGAACCC
1201 CACTATCATG CGGCAGGTGT TCCTGGGAAA CCCAGACAAG TGCCCAGTGC
1251 AGCAGGCCAG GTAGGCAGGT CCTGGCTGTC CCCACACTGG AGATGCCCTC
1301 ACCTCCTGGT CTGGCCCACA TGCAGTGGTG ATGCCTCAGG GTCTTTGTGA
1351 CTTGGTCTAT CCATGTGTCC AAGTCTGTAA AGGAGGACTT CTGCCAGAAC
1401 GTCCCCTTCC AGAGGCTGGA GCCATGACTC CCCTGTTACC CAACTTCAAG
1451 GTGCCTGGCA GGAACTTCTA TGATACCAGG CAGCCACAGA GGGGAGGGAT
1501 CAAAGTTGGG ACAGAGGCTG GTGTTTGAGA GACAGGATAG CCTAGACTGT
1551 GAACATGGGC AGTGGTTAGG GATGTAGACA TATGTGGTCA AACTGTAACA
1601 GAAAGCAAGG AAAAGGTACA AGCAACTCAG TTACCTTTAG GGGAAGAAGA
1651 GAATTAGGAG GGACACAGGG AGCTTCAAAC TGGGAGTGTT TTGTTTCTTA
1701 AACTGGGCCA TAAGTACATG GATGTGTGTT TTATTATTCT TTATATCTTA
1751 CACATCTATT TACTCAGCAA ATCTTACAGA ACTTCCTGTG TACCAGGCAT
1801 TGTTTCAAGT GCTTTAGAAA TCTCTCTCTT AAGTAGATGT GATGGGTGTG
1851 AAATAATTCA TGATGAAACC AAAGGGGACA CAGTAGGGCA CTCATGTGAA
1901 AGAAGGAGAG GTCTAAGGCA TAGCATCAGA GGCCCCAAAA TATCAGCTCC
1951 AACACCAGAG GATGCATTTT CTTTTTAATT AAACACTAAA TTTTCACTGC
2001 CCAAATTCAT TTGCTCAGCT GAATAATCGG TTGCAGGCCC AGCACCTGCA
2051 GTCCAACACT TGTGCTCTGT TGGTATGAGA GGGTGCTCAT TCCCACGCTG
2101 GCTCCCTCCC TCGGGCCATC TCCAGTGCCC CTGCCAGGCC TGAAGCCTGC
2151 CCCTGAGCAT GTGCGCCAGA GCCTCAAGGC TTGAGTGCTC CTAAACCAGG
2201 GCGGGAGGGA GCCTCTCCAC CCCTCCCCTG AACCTGGGCA ATCAGAACCA
2251 GCCCCTGATG GAAGCCTGAG CTCTGGGGCC TCCTGCCTCC CCCTCTTTGT
2301 GCAGCGTTTT GTGTAACTGC GTTCTGACCC TGCGGGAGAA CTCCCAAGAG
2351 CTAGCCAGGC TGGAGGCCTT CAGCCGAGCC TACCGGGTAA GACCAAGAAG
2401 GGCACCATGC TGTGTCCTCT CCCCTACGTT CACTCTAACA CACAGCCCAG
2451 AGCCCCTAGA GGAGGCACAC AGGGAAGGAA AAGCTGGTCA GGGATTGTGG
2501 GGAGACGGGG AGCAGCCTGG GTGCCTTCCT CTGTCTCACG TGACTGTGGT
2551 GTCTCAGGTG CCCTGGTTGG AATCATCCCA GTAGGATCCA GGTGGAAAAG
2601 CCCTCATGGC CCAGCTACCG TTGAGGGCTT AACCCAACT CCTGGCCCGT
2651 AGCCCTGGAT GCCTCATGAG ACCACCTTTC CCTCCCCCAC TCCCACTCCA
2701 AAGGCAGGTG CCGAGCCTCT GGAGGTTCTT CCCAGGTTTT TATCCCTTTT
2751 GGGACTTCCT GCCTAGCCCT TCAGAGAGAG TAGTCTACTT ACAATCAAAA
2801 CAAAAAGGTG ACCCAACCTG TTTCAAATT CTCTGGAAAG GGACTTGCCC
2851 TCAGGTGATT TGTGTTCTCA AGGGAAAGGC TGAGTCGGCC CCTCCATCCA
2901 GGGAGATGGA CTGCCCACCA CCCCTACTCT TGCCTCACTG GGTCCTGGGC
2951 CCACCCAGGG CCTGGGCTGA AGACCCTGTG CATGTGTCCC CAGAGCAGCA
3001 TGCGCGAGCT GGTGGGGTCA GGCCGCTATG ACACGCAGGA GGACTTCTCT
3051 GTGGTGCTGC AGCCCTTCTT CCAGAACATC CAGCTCCCTG TCCTGGTGGT
3101 ATGTCCCCTG CCCTCGCCCA TGGTACTCTT TTAGAGGAAG AAATGCAAGG
```

FIGURE 3, page 1 of 33

```
3151 CAGAATTGCC AGTTGCTTCC ACGAGCATGT GCATAAAATG GGAAAGACAC
3201 AGCTCTCCAG ACGCTGNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3751 TTGTGGGATC GCTCTGATCT CTCTGTTAAG TGAATGGGCC CTGTGGTGGC
3801 TGGTGACCTG GAGCACTCCA GGGGAAGGAA GGTGTTGAGT GGTCAGTCCC
3851 AGGGCCAAGT CCGCTGGTGG TGGCTCCCTC TGAACCAATA GGATCTTGAG
3901 GGGGTATATT GGTCTCTTTC AGGATGGGCT CCCAGATACG TCCTTCTTTG
3951 CCCCAGACTG CATCCACCCA AATCAGAAAT TCCACTCCCA GCTGGCCAGA
4001 GCCCTTTGGA CCAATATGGT AAAATAAGTG GGGTGTTCCT TGTTCTCTGG
4051 GGTTCTAGTC TAGGGCAGGG CACCAGCCCC TATAGAATGG AGTCTTGCAA
4101 GTGAGGCTGA GGGGGCAGTG GCTGGTACAT CTATAAACGT CTATGCAGTT
4151 GGAAATGCGG AGTCCTTAAG AGTCTGCTCA GCCTGGGCTC AACTGCACCC
4201 TCTCCTCAGA GCTTTGAACT CTGAGGAGGG ACCTCTCTAC AGAAATGCAA
4251 GCCCAAAACC CCATATTCAT TCCACTTTCC CTATGTGCCG GCCACCATGT
4301 TAGGCAGTTT AAGCCACGTT ATCTCATTTA AGGCTCTGCA TATCCCTGCT
4351 AGAGAAGCAT GACGAGTCTC CAAGGAAGCT ACTCCAGAG AAGCAAAGCG
4401 ACTGGCCCAA AACCCCACGG CTGGCAACTG GCAGAGCCAG AAGTGGGAGC
4451 CAAGCCCCCT GAAATCGAGT TCTGAGCTTT CCCCACTGCA GGATTCTGCC
4501 AGGGAATGTT CACTTCCATG GAAACAAACT ACTACACCCG TGTCTCTCTT
4551 TTCTTCCCTG ATCAGCTTGA ACCACTTGGA AGCAAAACAG AGACCCTGGA
4601 CCTGAGAGCA GAGATGCCCA TCACCTGTCC CACTCAGGTA GTAGGGGAGG
4651 ACCTGCCTGG CTCCTCTCCA CAAACCAGGG CACACAGCTC GCCCTACCCA
4701 CTTCGTCCTC CACCACAGCT TCCTCAGTAC CCATCTTGCC CCCTTACTGA
4751 GGCCTGAGAG ATTTGGAGGA TGGAGGGGAG TCCATGAGGA TGGACAGGGG
4801 AGGTGAGAGG GGAGACAAGA GTGCAGCTGT CATTGGGAAC AGGAGATGCA
4851 GCAGGGAGAG GAGGCCTGGG CCCCAGCAGA GGGAGAGGAT CCCGGTGAGA
4901 AAAGTGGGCT CCTGAGAGAG GAAATCAGGA TGCCAGGAAA ATGGCAGGAG
4951 GGCTTCTCTT AGCAGTGGTG TTTGGGGCAG ATGAAAAAAT CTGACTGCAG
5001 GTTAGAGGGC CCAGGCAGGA GCCAGGCAGG CTTAAGAGCT GTGGTTGGAG
5051 AGAGGAGAGC CTGGATTAGG GAGATTCCAC AAGGAAAGGA TCACAGAGGA
5101 CAGCAGCAAA GGGCAGAGCC CAGAGCTGTA TGGAGGAGGG ACGAGGGTGG
5151 GCCTACCAGG ACACGGCAGC TCCAGGCTCC TTTTAAGGAG GAATCCGTAA
5201 GTGGTTGTTA AGCTTGACTT CAGGCCTGGG GTGGGGGCAG GTTCTCATTG
5251 TCTTCAGCTC CTGTTTCTAG GCCCGGTCTT ATGGCTTTTT AACCAAATAA
5301 GGCCAAGGCC AGAAAACCCT CAGCAGCAAT AAAAGCAGAA GGCCTGACCC
5351 AATCTGGGAG GCTGGGTTTC CCTCCTAGGT CGGCCACACC ACCCTCTCCC
5401 ACCCTCCCTG CTGGGGAATG GACCTGCAGC TCCCCCATGT GTCTGCTGGG
5451 AATCCTGAGA GAGTGGGCAC CCCTGTTCAC ATGCCTGCTC CCTGTCTGCT
5501 GCCTGCCCTA CCCCAGTCTT GGGCTCAGGC TCAGTCTTGT GTGCCATCAG
5551 CCCCATCAGG AGAGCAAGAA TGGCAGGAAG AAGGGATGGG AAGTGAAGAC
5601 AGTCGTAGCA GAGGGCTCAG TTGCTGGGTC TTGTGCTTGG AGCTAAGGAG
5651 ATTGTCAGAT TCTGCAACAG CTAGTGCAAC ACAGATGCCT CTAGTCCAGG
5701 TGGTCAGGTG CTGGCCAAAG GCCTGGAGCA AAACCTTAGA GGCCCCTACT
5751 GTGCCAGGTG TAAACTCTTT AACTGCTTTC CTAAGGATGC CTTGGGGGTT
5801 CTAGGGGAGC AGCCAGGGAC CGTGGATAGT GGGGGCATTT GGGGACTCAG
5851 AAATAGCCAT ATTGTAGATA TTTCAATATT TTACCAACCC TATAGCCATA
5901 CTGAATATCA GCCATGGAGG GCCCTTTCCA AACTGTCCAC TCCCCTTCCA
5951 TTACATAACA AAAGCAGCCA TCATTTGCTC TTTCTTTCAA CAAACCTGTA
6001 TTGAGTACTG AGTTGGAGCC TAAGCACTGG GTCAGGGAGA GCCCTGTCAC
6051 CCTGGGCTTC GAGGCAACCA CTTCCAGGCT TTACCCCAGA TCAGGCAGAG
6101 ACCCCCAAAA GGAGGCTGCT CCACCCAGCA GCATCTTAAG CTGAGTGGGC
6151 TCAGTGCCTC CCTTCTAGAC AGAGCCCAAT GGAGCCACTG CACTGATTTG
6201 CAGAGGTGAG CAGATCCAGC CTCGTGGAAC CAGTAGAAGC CCAGCCCTGG
6251 TGAAGCTGTT GCTAAGCAAC ATTGGAGCCC ATTCTGAAAG GGTCCATCTG
```

FIGURE 3, page 2 of 33

```
6301 TTGGCCAGCC CAACTTCACT GTGTTCTGAG CATTCTGCAT TCCTCAGTCC
6351 CATCTGCTCC CTCCCATGTG CCTTGGAGTG ATATAAAAGT CCACCAGCAT
6401 CTCAGTGTGA GCTGACAGGG GCCAGGCAGC ACCTATTTTT GTCCTAGATG
6451 TGTCTAAACA TAGAGGCAAC AGGCAACAGG CAAGACGCAG TGGGGGGCGG
6501 GAGGCAGGAG GCCGAGATGG CTGTGAGCAT GAGCTTTCTC AGCCTCCTCC
6551 CTTCTCCCAT CCGCAGTCTA ACTGCTCATA CGTTCTGTGT GCCAGGTAGG
6601 GTGACTTAAC AGCACGCCAT GGATTTCTGT TGTAGTTTCA AGTTGGACAA
6651 ATTCTTTTAC AGACAACTTT TGACTAGCCT TCTGTGGACT GAGCCTATAC
6701 TCTGCCTTAA TGGGCTCTCT GCCCACTCCT TTCCTAACCC CAGGGCAGCT
6751 GGCTGAACAC CTGGTCCTTT TCTTAGGTTT CATTCTTTTT GACCTCTCTG
6801 AAGCCCTTGT CAAAAGTCAC CACCTCCCCC TTGAAATTCA CTCCTTCCTG
6851 GGTTTGTGGA CACTAAATCG CCTTGATTTT TCTGGTCTTC TGTTTGCTTG
6901 CCTTTAATGA CCCTCCTCCT CCCTTTCCCC AGTCTTGAAA ATGTAGATAT
6951 TCTCCAATTT TCATGTCTCC ATTCTATTTT CTTTCCTTTT TCACTCACTT
7001 TTTGAAACAG GGTCTTGCTC CGTCTCCCAG GCTGGAAGTG CAGTGGCGCA
7051 ATCACAGCTC TCTGCAGCTT TCAACTCCTA GGCTCAAGCC ATCCTCCCAC
7101 CTCAGCTTCC TGAGTAGTTG GGACTGCAGG CATGCACCAC CATGCCCAGC
7151 TAATTTTGGT TTATTTGTTT TGGTAGAGGG GGGGTCTTGC CATTTTGGCT
7201 CAGGCTGATT TTGAACTCTG GNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
7951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
```

```
 9451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9551 NNNNNNNNNN NNNNNNNNNN NNNNCAATAA ACCCAGCTAA AAACAAGCCC
 9601 AATAAAACCC AATAAAACCC ATTAGACAGG AACATAGGAG TTGGAAAAAA
 9651 AAGAAAAGAA GGGGAGGGGG AAGAAAGCCC TGAGGCACCC CGGCTGCCTG
 9701 TCTGCCACAA CCCTGGGCTG TAATTGTTCT TGCCATGGCC TCAGTCTGCA
 9751 ACACATTCTA GTGTCTCCTT GACCTCTAGC CCTCTAGCTC TGCCTCCCTT
 9801 TCCCCAACCT GTAGATCTTG TGATCAAATA GATTCAATGA AACACATTGT
 9851 CCAGTTGCAC TCGCAGCACT TCCAAAAAGG TCAAGTTTGT CCTTCCCTCA
 9901 GTGCCTCCCA TTCTGGTCAC GGTAGGACTG ACTCCAGCCC CTGGACCCTA
 9951 AGCTGAGTCT GGGCTCCTTT GACGTGCAGG GAGAATGCCA CTGAGTCTTG
10001 TCTCTGAGGA CCCTACCTCT CCAAATCTTG CCTCAGTTCC TCAGCAGGTA
10051 CTACACTGAC TGGCCATGCC ATTCTCTGAT GCTTCACTGC CTCAGCTTCT
10101 CAAGTCTGTC TCCCCACCTG AGCCAATTGT GAGTTTCTCT CTCTCCTCCT
10151 CTCATCCTGG CACCTAGAAA TGCTCTCTAA CGCTTGAGCT GCTCAACCAG
10201 CATGGGTCAC TTGTTTATAG CATGCTCCCA GATCGCCCTC TTTGTTGGTG
10251 AATGCTCAGG GAATGCTTAC TGTTAACCCG AGACAAGCCC AAGTAGCTAC
10301 ATGGACCTGC CACCATAAGC CCTCTCCTGT CTTATGCTGT TGTAGAGGGT
10351 CCAGGGCTCA CTTCTCCCAC TTGGCCCTGA GTACCTCTCC TTGAAAGGAT
10401 GTCAGGGGCT GGGCGCAGTG GCTCACGTCT GTAACCCCAG CACTTTGGGA
10451 GGCTGAGGCG GCGGATCAC CAGGTCAGGA GATCGAGACC ATCCTGGCTA
10501 ACATGGTGAA CCCCCCGTCT CTACTAAAAA TACAAAAAAT AAAAATAGCC
10551 ATTTGTGGTG GCAGGTGCCT GTAGTCCCAG CTACTCGGGA GGCTGAGGCA
10601 GGAAAATGGC ATGAACCCAG AAGGCAGAGC TTGCAGTGAG CCGAGATCGC
10651 GCCACTGCAC TCCAGCCTGG GCAACAGAGC AAGACTCCGT CTCAAAAAAG
10701 CAAGCAAGAA AGAAAGGATA TCGGTTACCT GTTTCAGACA GGAATGCTGA
10751 GACCAGGGAA AGGGGAGACT TGTCGGGTGC CTCAGGGAAC CAGTATCTGA
10801 GCTGGGGGCT GAGAGCTCTG TGTGGGTGGA CTCTGTCCTC CCAGTCGCTG
10851 CTGAGTCCCT CTCTTCCTTT CCCGCTGTCT GACCAACAGG GTTTTTGTTG
10901 GCCTGACCTC CAGTGTGAGG AACGGAACCA GGCAAGAGGC TTGTCCAGTC
10951 AGCTCTGGCC CCAGTTTGGC ATTCATCACT TGTTCCCTAA CCTGGAACCC
11001 GTCCCTTCCT CACTCTGGGG CTCAGCAGCT GCTCATCTAT AAAGTGGGGC
11051 ATTTGGGGGT TGCAAAGTCA GTCATCTACA ATGCCAGGCA AGAACATGGT
11101 TGCGTACATG TGGTCAGGTA TGAGACGAGA TCACTTTTCC AACATTCTGG
11151 TTTTCCCTTT TTTTTTTTTT TTTGAAACAG AGTCTTGCTC TGTTACCCAG
11201 ACTGGAGTGA ACTGGCATGA TCTCATCTTA CTGCAACCCC TGCCTCCCAG
11251 GTTCAAACAA TTCTCATGCC TCAACCTCTC AAGTAGCTGA GATTACAGAT
11301 GTGTACCACA CCTGGATTTT TTTTTTTTTT TTTTTGTAT TTTTAGTAGA
11351 GACAGGGTTT CACCATGTTG TCCAGGCTGG TCTTGAACTC TTGGCCTCAT
11401 GTGATGAACC CCCCCTTGGC CTCCCAAAGT GCTGGGGTTA CAGGTGTGAG
11451 CCACTGTGCC CAGACTAATT TTTTTTTTTT TATTGAAACA GAGTCTCACT
11501 CTGTTGCCCA GGCTGGAGTG CACTGGCACA ATTATAGCTC ACTGTAACCA
11551 CAAACTCCTG GGTTCAAGCA ATCCTTCTGC TTCAGCCTCT TGAGTAGCTA
11601 GGATTACAGG CACATGCCAC CATGTTGAGC TAATTTATTT TTTTAATTTG
11651 TCGTAGAGAT GGGGTCTTGC TATGCTTCCC AAGCTGGTAT TAAACTCTTG
11701 GCCCCAAGCA ATCCTCCTAC CCTGGCCTCC CAAAACGCTG GGAATACAGG
11751 CATGGGCCAC TGTGCCAGCC TGGTTTTTTC TTCTTGTTCC CATTTTATTC
11801 TCACATTTTC AGACCATGGG CTTACTACTC CACTGAGCAC ATTTTGTGAG
11851 AGTGCTCACA GCCCTGGGCC CGGTTGCTGT TTCCTGATCT CAGTCTTATC
11901 AACTTGATCT TGCTTTGCTG TCATTTATAC ATTTTCTCAT TAGCTTTCTC
11951 CCCATTTCTT CTTTGTCTGC TTCCTTCTTC CTTCTTTAAC TAACTCCTCA
12001 CCTGCAACTG GGGGGACTTG GATTCTTGAC TGGGCTTGTG TGAAAACTGA
12051 TTGTAAAACA GATAGGTAAG TAGGGAATGA GGAGGGTGTT TTACAAGAAA
12101 AAAAAAATGA CTAAGATACA GGAACCCAAC CTAAAGAGGA AAAGACATAC
12151 AGTTCAAAGG AGGCAGAAAG AAAAACATTA CAGATACTCA AATATATTGA
12201 TAATCATAAC ACTTTCTGGA AGATTAAAAA AATGCTGAAA CATGAATCCC
12251 TTGCTAGAGA AATTACAAAG CCAAGAAAAT AGATAGGTCT GAGGATTAGG
12301 GAGCTGTTCA GTTGCTAGGA GGAACACAAA AGCACAGACC CCAGACTACA
12351 ATGGGTATGA AACCCTCTGC ACGCCTTTTG TTGTCCATCC CTTGCCAAAG
12401 CTGTTATGTA AAACCCTCCG GGGGAATGAA TGAAATTATG TTTATACAGT
12451 TCTTTCTATA TAAGTGCAGA AGAATCATGT TAAATAAATC TACAGGGCAG
12501 GATTGTTAGT TTTTCTCCTT CTCAAGCAAA CTTCAGTGCT GTCAGATAAC
12551 TTCTCCATGT GTTTTTTTTT TCTCTTAGAA TGAGCCCTTC CTGAGAACCC
```

```
12601 CTCGGAATAG TAACTACACG TACCCCATCA AGCCAGCCAT TGAGGTAACC
12651 CCTGACTCAC ATCTGCCTCT CTCAGACACA AACCATTTCC ACCTGCCAGG
12701 GGCTCGGGTG TGGTACAGGT TTCAGAGTAT TCACTGAAGC AGAAATGTAC
12751 TTCTTACATA CTGGGGATTG GAATGTACAG AAAAGGCTCC CGGACCACGA
12801 AGCCCCAGGA TTGTCCTAAC ATGTTCTCAA GTTGCTTACC TGACGTCAGC
12851 CCCCAAGCAG AGGAAGTGTC TATGGATCGA TTTTCTTTGA CCTTGGCAAT
12901 CCTGGGCTCA CAGACGTGGT TACTGCTTAG GCAGCTCAGC CTCTCAAGAG
12951 GGAGAGGCAG CTGGTGTGAT GTGGCGTTGA CTTCTTGGAA GGTGGAGGCT
13001 GAGTGGGAGG GAACTACAAT TCTGGGGATG GGACCCAAAA GGAAGTGGAG
13051 GCACGTTGTT CATGTTCCTG TGGGCCCCTA GGCCTTGTTT GGTTCAAGTC
13101 AATCATTCTA GTGCTGAGGA TTCAGAGCCC ATGGTTAATT CCATTGGATT
13151 AACCATGTCT GTGAGCCTAG GACGGCCACT GCAAAGACGG CCTGGAGGAC
13201 CCCGGACTAT ACCATGACTG GCAGTCAGGC CTGGTCCGGA TCAGGTCTGT
13251 TGGTCACCAG GATGGGGTTT GACCCGCAGT TTCAGTTTCA CACCTATATT
13301 ATATCCAGTC TCATGTTAGG GGCTAGAAGG CATGCAGAGA AGTATCGAAC
13351 ATGGTCCGGA CCAAGGGAAG TGAGAGCCCA GTAGAATTTC ACAATTATTG
13401 AGCACATACT ATGTGCCAGA CACTATTCCA GGAAGACAGA AATGTTAACC
13451 AGACAGATGG ATCCCGGCCC TCACGTAGCT TACAATCTAC TGAGAAAGGT
13501 GTCTTATATA CATGGCTAGG CATGGTCATT TCAGATAGTG ATGACAGCTC
13551 TGAGGAGCGT GATGGGGCTG GGGCAAGGGA GGCAAATTCA GGTGCACCAT
13601 GCAGGCCAGG CCTTCCTGAG GTGAGATTTA AACTGAGACA TGCATAATGA
13651 GGAGACACTT GCTATACAGG GAGCCAGGAA CACAGTCCCA GGCAGAAGGA
13701 CCATGGACCA CACAGGCTCA GAAGTGGGAC TGTGTTGGGT GTATTTGGGG
13751 AAGAGAAAGA AGGTCAGAGT GGCTGGGGGC ATGAGAATGA GGTGGAGAGT
13801 GGGGGAAATG AGATCAGGAG TGCCAAGGAG CCAGATCACA CAAAGCCTGA
13851 ATTACTGAGT AAAACCACTG GATTTCAAGT GGAGAAAGAT GGGAAGGCAT
13901 TGGCGGTCTC AGGAGAGAGT GACATGATCT GGTTCACGTC TTTCAAAGAT
13951 CTCCCTGACT GCTATGTGTA GAATGGGTTG GCCATCAGCA GGAGTGATTG
14001 GGGAAAGACA TTTTATAAGC CAGCTGAAGA AACTAACCCA TATGAAATCA
14051 TTAAGAACTA TTGGATGCTA AGCTCTGGGG TGCAAGCAAT ACCAGATTGC
14101 TGGCTGCGGG TTATGCTGTG TCCAGCCTCT CTGAATTTTC TCAGGCTCAC
14151 GTTAGCCCAG TGGAGGCTTG TCCTCATTGA ACCAGTGACC AAATTCCCTG
14201 AGAATTGAAA CGTCAGCTGC ATCTTGTGAA TCAGGCATTT CTTCATTTAT
14251 TCATTTACCT ATTGGATGCC TATGTAGAGT GGGCACTGCA CTAAGTGCTC
14301 GGTAGACAGT GGTGAGCCGA ATGGGTCTGG ATCTGCCCTC TTGGTTCTTC
14351 AGTCTCATGC ATCTTTGCTT TTGCTGCTGG AAGAGCTAAA AATCCCAGAG
14401 CTAGAAGGGC GTGTGTTTGT TTTAACAGCT TTCTACTCAA AGTAACCACA
14451 GAAACAAAAT TCTGTCATCT GAGGTAACGT GAATGAGCCT AGAGGACATT
14501 ACGTTAAGTG AAATAAGTCA GGCACAGAAA GACAAATACT ACATGTTCTC
14551 ACCATATGCG GAAGCTTAAG AAGTTGACTT CACAGAAGTA GAGTATAAAT
14601 AGTGGTTATT AGAGGCTGGG AAGGGTGGAT GGTGGTTGGG GAGTAGAGAT
14651 AGCAGAAATT GATTAACAGA AAATTACAGC TATATAGGAA GGAGAATTTC
14701 TAGTGTTTTA TAGCACAGTA GGGTGACTAT AGTTAACAGT TTACCATATA
14751 TTTTCAAATA GCTAGAACAG CAGATTTTGA ATGTTCCCAA CACAAAGAAA
14801 TGGTAAATAT TTGAAGTGAG GGATAGGCTA ATTACCCTGA TTTGATCACT
14851 GCACATTGTA GAGATGTATC AAAATATCAC ACTATGCCTC ATAAGTATGT
14901 ACACTTAATA TGTCAATTAA AAATAATAAA AGCAAAACTA ATAAAGTGGC
14951 CACAAAGAGG CTTTACCTGG GAGCTTTTTA GAAATGCAGA GTCCTGGGCA
15001 CCACCCCAAA CCTGCTGAAT CAGAATCTGC AGCTTAAGAT CTTCAGGGGA
15051 TTTGGATGCA CTGATTTTGG GTGTGGTGCA TGGTTCTTCC CTTGTGACGG
15101 ATGAGCACGT TTCAATTCCA ACCAGGATCT GTTAATCTAC ATGGAATATG
15151 CTTATCTCTG GTTCACCAAC TATCTGAGAT ATATCTCATG TGCTGATGGC
15201 TGAATAACTT TTTACGTTGC ATTTTCTGTG AGTATTTGTC ATCTGCACAC
15251 AAGCATGCTC TTGAGTTCAT TAAACCTTTA AACAGAAGAA ATCCATCAGA
15301 ATGATGAATT GAGCAATCCC TTGGGAAAAA ACCAAATTCC ATAGGATTAA
15351 GCAAATAATA TTTAAAAGAA GTTCCATTTT TGCTCTCTCA TGATAGGAAT
15401 ATTTCAACAA GTCTTATCTT CATCATCTGA CTGAACAGAT GAGATGAGTT
15451 TTCATAGCAT CTGGCAGTCA GACTCCTGGA CAGTCAATCT GCTGGTCAAG
15501 CCCTACTCCA TACTCAGTAT GCATATATTT GAGACTTTGG GAAGATACTC
15551 AATTTTCCCC CAGATTTCTG GTACTAATCA TTTCTATGCC CTCTGCTTCC
15601 CATCCCACTC CTTTCCCCAG CACCTGGAAA ATATGTTCTG TATTAGAGAC
15651 AAAGAAAATT GACTAAAAGC ATCCAGGGTT GCTTACATCA ATTTAAAAAC
15701 ATATAAGGAA TAAGGCTGTT AAGTTAAATA TGCAAAAAGA CATACAGGTA
```

FIGURE 3, page 5 of 33

```
15751 TCCAGAAAAG ACAGGCAGAA ACCAGGAGCT TTACAATTTT AAAATATTTT
15801 GTGTTATTAT TCTAAAAATA TTTTAATTAT TGTCTAGGTT CTACCATTAT
15851 AATTAGTGTC AGTTAGCTTA ATTTTATAAA ACACACATAC CTGTAATCTC
15901 ATGTTAGGCA TCCAAATGCT GTGTTCCTTT GGGAGACCCA CCTGTGTAGG
15951 ACTTCATGGT TTTCTTCCCT GCTTTGGGGC AGCCACTGGC TCCATTCAAA
16001 GCATAGATAT ATGGGGATAA GAAAGGTTGT GTGTGGGTGC ACATGTGGAG
16051 ACATGCACTA TGGGTTGTGC ATAGGGGTAG CTAGACACAC CCATTTCTCC
16101 CCCTTTAATT TCCCTCCTAG CCCACCTATA ACTCACAGTT CTTTCCCTCA
16151 CATGATCCTG TATGGTGACT CATTTCTAGC CTCCATCAAA AATCCCTTAG
16201 CTGGTTCTTC TTGGGCTGAA GCTTATCTCC CTGCACAATG AGTGTTGGGC
16251 ACTGAATCTT TTCTCCTGTT GATTTAGAAC TGGGGCAGTG ACTTCCTGTG
16301 TACAGAGTGG AAGGCTTCCA ATAGTGTTCC AACCTCTGGT GAGTGAAAAC
16351 ATCATCATCT CCTTCAATTA AGGGCCTTGC CGAATATCAG GTTGTGGGGA
16401 GACCCTGCAA ACATACCCTG GAGCTTTAAG CAGGACTTGC TAATTCCCCT
16451 GCAGTGCAGA CCTAGATCCT GCGGCCTGCC GCCACAGCTG GGCTTCCATG
16501 TGGAGGTGCA CAGAGCTCTC CATTGGATGC TACTTCTTGT CTCCTTATAG
16551 TCCCAGTGGC AGTCCCTTAG GCCTCCCTGC CCAGTGAGGC AGGTAGAGTC
16601 AGGGATTGGG ATCTACCTGC CTGTGCTACA TGACCCTGCA GCTGGAACTT
16651 TCCTGGACCA CCCCAATGTC AATCAGGCTC TTCTGAGGGT GGATGATAGC
16701 CATGAAACCC ATTCCCTGCA GTGCCTTGGT TGGTCTGAAT GAATGGGAGG
16751 GGCAAAACTG CTAAAGCCTT AAGCTGAAAA TAAGTACAAT GGGGAGCAGT
16801 GGGACAGAGT TATAGACTTC TGGTAAAATG TGTACTTTAA GAGGTAGATA
16851 CCCCCAGCCC CCACAACCAC CTCTCTGCTT GTCTCCCCTA GTCCACCAGC
16901 TCCGACCAGC AGACATCAAA GTGGTGGCCG CCCTGGGTGA CTCTCTGACT
16951 GTGAGTAGTG AGCCATGAAC CAGGATGGGC AGCTCAGAGT CCAGCCAGGC
17001 CCTGCGCAGA ATCTGTGCTT CCCCAGCATT GGCTCCGCTT TCAGTGCTGA
17051 GCCCGTGTTA CTGAGGGCCT ACCCATGTCA GGCACTGAAA CACAGCCAGG
17101 AGATGTAGAA TGCCCTGTCT CGCCACCTTC CCAGTTCTGC TCAAAGCCCC
17151 CTCGTCCATG AGGCCTCCCC TCAATTCCCC AGGGAGAAGC AATCCCCGCC
17201 TTCCCCACTG TTCACAGGCG TTTTGTTGGT GTGTGATGGC ACTCATTCAA
17251 GTCTGCCTGC CTTCATCAGG GGACTCATCT CCATCTACCC AGACTCAGAG
17301 TGGCAGGTCT TACACACACA CTGCCCCATG CTCCCTACTC CATTTAAGGA
17351 CATGTGCTTT GGGGCAGAGG GAGCCCGGTT CCTCACACAT AGCACAGTCT
17401 TGCTAAGTGA ATTGTGTTCG CCAATTACTT AGCCATTGTT GTGTACACCA
17451 ACACTCTATT AGCAATTCTA AGGGAAATGA GGTATGAAAC ACAGTCATAG
17501 CCCCCCAGCA ACCTGTCTGG CTGGAAAAAC AAGAAACGTA CACAGAAAGA
17551 AATGCATAGT CACATAGATG ACATATAGGA CTTGGATGTT TTATTTTTAT
17601 TTTTTAACTT CTAAGTTCAG GGGTACATGT GCAGGTTTGT TACACAGGTA
17651 AACTTGTGGC ATGGNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18801 NNNNNNNNNN NNNNNNNNCT CCAAGTTTAA TATGGCCTGG AGAAAAGCCG
18851 GTACTATTTT TAGAAAAGGC AAATCCAGGT CCTAGCTGCT ACCCCAGGGC
```

FIGURE 3, page 6 of 33

```
18901 CAGAGGAAGG CTCTTCCAGT GCCCTCAGCC TATACCCCAG CCCTGAACTT
18951 TCTTTTTGCT TTTTACAGAC AGCAGTGGGA GCTCGACCAA ACAACTCCAG
19001 TGACCTACCC ACATCTTGGA GGGGACTCTC TTGGAGGTGA GGATGTTCTT
19051 GATGCATGCT CTATTGATGA TGCTCTCTCA GAGAGGTGTG AGTAGTGTGT
19101 TTCCTGTCAC CCCTCCAGGG ATGCAGTTGG GTCCCCAGGT CCCAGCGCTG
19151 AGACAGGAGA CTCAATGCTT GCATTACCCC TGAGGGTGAT GGGAGAGACG
19201 CCCCAGGGGC CCAGAACCCG GTTCCGGTTC TGGCTTGTGC ATATGTTGAC
19251 ACAGGGAGCA GCATGTTGGT GTGAGTTTAA CAAATATGCT TTCTCCTCCC
19301 CAGCATTGGA GGGGATGGGA ACTTGGAGAC TCACACCACA CTGCCCAGTA
19351 AGTAGCAGCC CAGAGAGGCA CCATCACTGT GGCCGTCCTC CCTGGGGCCA
19401 GGGCCTTCCT GCTGGAGGAG GGGAAGAGGA GGTTATCTGC AAGAAGGGAA
19451 GTCAGCCAGC CCTGAAAAGC CCCAGACTTC CTGTGTCCCA CCCATGTCCC
19501 CACCCTGCAT GCTCATCTCA GTTACTGTGA GGGTCCTGCA GGCTCTCACC
19551 TGTGCTCTTC TCCTCCTCCT CCTCCTCTAA AGACATTCTG AAGAAGTTCA
19601 ACCCTTACCT CCTTGGCTTC TCTACCAGCA CCTGGGAGGG GACAGCAGGA
19651 CTAAATGTGG CAGCGGAAGG GGCCAGAGCT AGGTGAGTAG ATGCCGTACA
19701 GGAGGGCGAG NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20951 NNNNNNNNNN NNNAAAATCA AAAATTAGCT GGGTGTGATG GTACACGCCT
21001 GTAATCCCAG CTACTCGGGA GGCTGAGGCA GGAGAATCAC TTGAACCCAG
21051 GAAGTGGAGG CTGCAGTGAG CCAAGATCGT GCCATTGCAC TCCAGCCCGG
21101 ATAACAAGAA TGAATCTCCA TCTCCAAAAA TAATAATAAT TAAAAATAAA
21151 TAAAAGATAC AAAGGAATCA AAAGATGAAC TCCCTGGCCA CGAAGAGCTT
21201 GCACTCTAGG TAAGGAGGCT AAACAAATGG GAAATAACTT TCTGAAAAAG
21251 ACAATGCTGG GTATGGCAAC AATGCAGTGC TTCGCATGGA GTACAATTAA
21301 GAGAACAGAA GAGCACACAG TATGAACTGC ACTGTCTAAA GACAGATGCA
21351 GACCCAGAAG GGACCCCTGA AATCATCCAG TCCAACCTCT TCCTTTAAAA
21401 GATGGGAAAA GTCAATCCTA GCAAGATTCA GCAACTTGTA CAAGCTCAAC
21451 AGCAAGTTGG TAGCAGAGCT GAAAGTAGAA CCACTGGTCC CTGGGGTAAA
21501 AAAGGAAATG CAAGATGTGT GGATCAGGGA GCCCAGAGAG GAGGCTCAAG
21551 GGAAAGTAGG ACTTGGTCTG GGCCTGAAGG ATGGGAAGAA GATGGCTAGG
21601 AAGAGGGGAA GAAGCGGCAT TTGTAACTTC CCCTCCTACC CACGAGGGCT
21651 TATTGCCCAT GGATTCTCTT AGTCACACCT TGAACCTGTT AAAAGGTTAA
21701 AGGCACTTCT GTGGTCACCT TTGACCAGAA AAGTCTTTCT TTATAGCTTT
21751 CTGGTATACT CATCAATAGC AATAATGTAT GGGATACAAT CCTAGATCTG
21801 TAAATTCTCC TTAATGAGAA ACAAGGGTAG GGATGGTACC ATGTGTGTTT
21851 GGCCACGTAC CTAGCTTACC GTGGCCACCC AAAGATTTTC AGTGGCCAGC
21901 TCGCACTGGT TGCTGCTTTT ATGGCTTCTT CCATGGACGC TTTTCATTGGC
21951 TATATCCCCT TTGCTGACCT AACTTCTCAC AGACATTTCT TTAAACACAG
22001 CTTTATTGAG GTATAAATGA CATGCAATAA ACGTCACATG TTAAAAGGAT
```

FIGURE 3, page 7 of 33

```
22051 ATAATTTGGC TGGGTGCGGT GGCTCACGCC TCTAATCCCA GCACTTTGGG
22101 AGGCCAAGGT AGGCAGATTG CCCAGGAGTT GGAGACCAGT CTGAGCAACT
22151 TGGTGAAACC CTGTCTCTAC CAAAAATACA AAAAATTAGC CGGGCATGGT
22201 GGCATGGACC TGTAGTCCCA GCTACTCGGA AGGCTGAGAT GGGAGGATCA
22251 CTTGAGCTCA GAGGGGGTTGA GGCTGCAGTG AGCCGTGATC ACACCACTAC
22301 ACTCCAGCCT GGGCAACAGA GCAAGACCCT GTCTCAAAAA GGATACAATT
22351 TAACATTGTA CCTGTGAAAT CATCACCACA ATCAAGATGA AAAATGTGTT
22401 TATCACCCAC AGGAGTTTTC TCAGGCCCCT TGGTAATCTC TCCCTCCTGC
22451 TCCTTCCTGT CCCTACCTCA CACCCCAGGC AACCACTAAC CTTCTTTCCA
22501 TCACAATAGA TTAGTTTGCA TTTTTAAAAA TTTTATATAA ATGGGATCAA
22551 AGAGTATATA CTTTTTATCT GACTTATTTA GCAAAATGAT TTTGCGATGC
22601 ATCCATGTTA TTCGGTATAC CAATAGTTCG TCCCTTTTTA TGGCTGAGTG
22651 TAGTGTTCCG TTGGCATTCA TATCGCTCAT CCAGAACACC AAATGGTATT
22701 GTTTTATTTA TGGCAGACAT CAGGGGATGA AGGGAGAACT AATCCTGTCC
22751 ATCCTGGTTT ATTGGAGAGG GAGAAAAAAA AAAGTGAGGA GATGGGGAAT
22801 GGTGCGGAAA TCTAAGTAAC CACAGAAAAG AAAAACAAAA GGATTAAAGG
22851 AGCAGAGAGC AGGGCTTAGA AGTAAAGGTT AAAGGAGTCA TTAAGCCTGG
22901 AAAGGAGAAA ACTGAGGGAT AATTGTGAGC TGTGACTTTT CTCAAATATA
22951 CAAAAGGTTA TTTTTAAAAC AGGCAACTGA AGAAGAAATG AACAGGCTTG
23001 GCTTACGAAG AAAGAGCTTG AGGAAGTATA AGGGAAAGTC CCTGAGGGGA
23051 GGCTTGACGG GATCCCAACC CGAGTGGCCG ATGAGACTAT TGGGTGGCAG
23101 GGGCTAGATC AATGTGGCTC CAGGGTCCAG GGCAGCCATG TGATTGTTAC
23151 TAAGCTGAGA TTTCTTGAGA ATGGAATGAC CTTTGTACTG GTAACATCAT
23201 TCTTCTTGAA ACACCTCTCT TCCTAGGCCA AAATCCCATG TCGTGAGTCC
23251 TCGCTCCTGA GCCGGCACTA ACGCCCCTCT CTCTACCCCC CACCTAGGGA
23301 CATGCCAGCC CAGGCCTGGG ACCTGGTAGA GCGAATGAAA AACAGCCCCG
23351 TGAGTACAGG CCCCCAGGCC ACCCCTGAAA GGTGCCCATC TCCTGCTGGC
23401 TGGGGAGGGG ACAGCCCCAT AAGGGTCCCT CTCACCACAG CACTTCCTGC
23451 TTTGGGCTAG CCAAAAGATC CTCGGAGAAG CAGTCCTTAC CAAGGAGGCG
23501 CCTGCCCTGG CCACACTCCT AGACGCAGGC TGTGGCACCC CTCACCCCAG
23551 GGCCGGCTGC GGGAGGGCAA GGTGGAACAG GGAGTTGGCT GAGGTGGTGG
23601 CCTTGGCCTC TGACAGCTTC CTGCTTTAAC CAAGAGGTGG CTTCCCAGAG
23651 CCCTATTATG TAAATGCAAG GTTCTAAAAA TAGGCTTCTC ATTCCAATCC
23701 AGTTCTGCCT CCTTCCCCTC ACCCTGCCCC TCTGAAACTT CTCACTAGCA
23751 CTTTTTTTTT TAACCGTTCA GTGTTTATGC CTAGGAATTC AGCTCCCGGT
23801 GGGATTCCTA TTATGGAGGT GGCCAAGTGG AAAGCCAACT GCTTAGAGGG
23851 CCTCCCAGCC CCAACCCCGC TTCTCAGTCC ACGCTGGGCT CTTCCTCCAG
23901 TCTCCTTCCC CCGACCCTAA GAACTCATCC CAGGGGCAGC TTAGGGCCTT
23951 TGCTTCTAGC TGCATCCTTT GCCTACAGCT CCCTGGAAGG CCTTCATTTG
24001 GGGGGACGTG GTAATCCCCT CGGCATTTAA TGGGCCAAGG ATATGTGGGA
24051 CACATCCACA TTCTACTTCT CCAGGGACAC AACTTTCTTA AGATTTCAAG
24101 GGGAAAATAG CCCTCCCTTG TGTAAGCAGA ACCCCGTCCC CCGCCAGCGC
24151 CCACCGCCAA AAAAAAAAAA CATCCTCTCT GTGGAGCACC TTATCCTAGC
24201 ACCAATTGAG GGCTGGGAAG CCCCACTTTG TTGCTTTTCT TTTTTTTTTT
24251 TTTTTAGACG GAGTCTCGCT CTGTCACCCA GGCTGGAGTG CAGTGGTGCG
24301 ATCTCAGCTC ACTGCAGCAT CTGCCTCCTG GGTTCAAGTG ATTCTCCTGC
24351 CTCAGCCTCC CAAGTAGCTG GGATTACAGG CACCTGCCAC CAGGCCAGGC
24401 TAATTTTTGT ATTTTTAGTG GAGACGGGGT TTCATCATGT TGGCCAGGCT
24451 GGTCTCGAAC TCCTGACCTC AGGTGATCCA CCTGCCTCAG CCTCCCAAAG
24501 TGCTGGGATT ACAGGTGTGA GCCACCGTGC CTGGCCCAAT TTGCTGCTTT
24551 TCTCTGTTAC AGTATAAATA AGACAAAAGG CACTTGGAGG GCGGGCGGGC
24601 TGGCCCAGTA GCAACATTTG TATGTGCCTC CCACCAAGGC CTAAACTCAG
24651 GATCTTCTGT CCCCTCAGGA CATCAACCTG GAGAAAGACT GGAAGCTGGT
24701 CACACTCTTC ATTGGGGTCA ACGACTTGTG TCATTACTGT GAGAATCCGG
24751 TAGGCCCCCG ACCAACCCCA TGGGGACCTG AGAAGGAAGG TGCTGACCTC
24801 TGGCAACACC CTTGCCCATC CATCCCTGGC CCTGCCCCGA GCTCCTCGCT
24851 CATGGGAACC ACATTTGCCT GCTGCCCCAG GCCCTCCCTG GTTTACACAT
24901 GCCAGGCAAG GCCCAGCCTT TTCTACTGCC TGAGCGACCC CTGGAAGAGC
24951 AGGTGCATTG GTTCCCCAAT TCCAGAAGTA AGGCCAAGGT GGACCCACTG
25001 TAGGCACTGC TGAGGTGAGG CCTCTCTTAT CCACACAAAT ATGACCTCTG
25051 GTACCAGATA GGGGACTAGC CATCCTCACC CCATCCCTGC CCTGTTTCAT
25101 TTTGAGGAAG GGCAAAACAA TGTTCTAAAT GGGGTTGGAT GGGTCATCAC
25151 GAGTTAACCA AACCTCAGTG GTGGCCCTGG GAGCCCAAAC CTGTTCCTGA
```

FIGURE 3, page 8 of 33

```
25201 TGTTTCCATG GGGAAATATG TCCAGACTCA CAAACTTCTG GAAGTTGTAA
25251 CTTCCAAAAG TTTTTATTTT GGAATACATC TTGTTCTAAG TTGGGGACAT
25301 TCTGTATTTT ATTTGGCCGG GAAAAGGCAG CTGCCCAGCC TCAGAGATTG
25351 TGGTCGAATG TTTGACAGTC ACAGATGGTG GTGTGGGGGA AAAGGCATTT
25401 ATCACTCCTT GTGGCTCACA GAGGCCAACT ACCAGAGGCT TGATGAGAAG
25451 TATCCTCCCA CACACAGGAG TTGGTTTTCC AAACCTCTTC CCCTCAGTTC
25501 TCTCCCTGAC CACCCACACT CTAGAGGCAG AAGTGACCCT AACTCATGGG
25551 GATTTAAGTA TTGCTGTGTT CTGGACTCCG GGGATACCTG GACCCCAGAG
25601 CTGCGTGGAC ACTGTGGACG CTGGCGTAGG GAAATGCCCT CTACTAGTCC
25651 CAGGATGTGT GTTCTTCATG GGAACAGTCA GGTTTATACT TCCGAGAGCG
25701 TAGTTTAGTT GAAAGGGCTG GGCTGCCCCG ACTAGGATTA ACTCAGACTG
25751 TTTTAAAAAG AAGGAAAGGG GGAAAGGCAG AAACTCTGGG AGACACGAGG
25801 TCCCCTCTCC TCTATTTTAA CTCTTCGGCA TGGATTGTCT ATCTTGTTCC
25851 TTTCCCCTTC TTCCCAACTC CCAGTACCCC TTCTGGTGGC TGTGGCCAGA
25901 AACTCAGCGA ACAGCACTTG TCCTATGCCC ATCAGTGTGC TGGAATTGAG
25951 TGCACGGTAT CTCACCTGGC CTCAGCTCTT CGTCTCCAGA AAAAATAATG
26001 GGCTGCCTGA GCTCTCCCCT CCTGCCTGAG TGGTGCTGCT TTGTGGGTGC
26051 CTCAACTCCC ACTTCCTGTG GACATGCTTT CTTCCATGAG TATAAGAGCT
26101 TCAGGTTACC ACCCGCACCC CCACTGGTAT CAGCCTGTGA CACCTTCTGG
26151 GCCTGTAGCC CAGAGCCACA TCTAAAAATA GAGGCCCATC TCCCTCTGCT
26201 ATAAAGCAAA GCCCTGAGAT TCAGCCTGCA AGGACTTACT GAGCACCTAC
26251 TATGTACCTT GTTTGCATCA CCCAGGATGC TGTGGACACA CCTCTAAATC
26301 AGCCTCCTAC TGGGGAGATG GTTCAGAGGA AGAGAACCTT ACACTGAGTC
26351 ACAGGGGATA GAAGTTAGGG GAACACAGGA GAGCAAAACA TTTCAGGCAG
26401 TGGGACCAGC ATGGACCAAA GCCCAAAGGA AAAAGGAAGT GTGGCCACCC
26451 AGGGCATGGC AAGGGGCTGG AGAAGGCTGA GGTCAGATGA CGGATGGGAC
26501 TGCCAAGAGC CAAGGCCAAA AAGTGGCAGG ACCCAGCACT GGCAGAGTCC
26551 ACTGTTGGGT CTGAGATTAT GTAGAGCAGG GTGGGGGTTG GGATTGTTCA
26601 TGGTGTCTAG TAGGGGACAA GGGATGATTC CTTACAGAGA CTCAGCAGCA
26651 ACAAGAACTG GGCTTCTCAG TTTGACCAGG ACCACCGAAG CCCCTCTGTA
26701 CCCACTCAGT CATTTAGCCC AGGCCCCAGA GCCCTCCTAT GCTCTTGCCA
26751 TTCTCTCAGA GCGGGCACCA GGGGCTAAAG AGAGTACCCT TTTTTCCTTA
26801 CAGGAGGCCC ACTTGGCCAC GGAATATGTT CAGCACATCC AACAGGCCCT
26851 GGACATCCTC TCTGAGGAGG TAGGAGAGGG GTTACGTGTT CCTGGGTCCC
26901 GCCAGCCACC TCCCTGGGAT GCATGTAGGC AGGCTGTGTT CAGTGAGATG
26951 CTCACGGAGC AGAGACCCGC CATGAGTGAG CACCTGGATG GCAGGGAGGG
27001 AGGTGGCTGT CAAGCTCCTC TGCAGGGAAA ATTCTCACTT GGCCAGAGAC
27051 AGGGTTGTGT GGTAGCAATG AGCTTGCCTC TGAACCAATT GGCCCAGGTT
27101 TGCGCCCAGC ACTGTGGCTT CGGGCAAGTG ACTTCCCTGT GTCTCAGTTT
27151 CTCAACCTAT AAAGCGGGGC CACTCAAGAA GATTCAGTGA GATACTACAA
27201 GTTGCATCCC CTCTCTGGGC CTCAGTTTCT TCATTGGTAA AATTGAGGGG
27251 AGTGGGAATT GGATTGTAGA TGACCCCCAG GTTCCTCCCA GCAGTAGCCA
27301 GTGCCCTAAC GAAACCACCC TCCACTCCCT GCAGCTCCCA AGGGCTTTCG
27351 TCAACGTGGT GGAGGTCATG GAGCTGGCTA GCCTGTACCA GGGCCAAGGC
27401 GGGAAATGTG CCATGCTGGC AGCTCAGTAA GTGGACAGGT CACCGTCCCA
27451 AGGCAAGGGC ACCTGGGGTG AGGAGGGCTT GCAGGTGCCA AAGGAGGAGA
27501 CCAGTTGAGG CAGAGCCAGG CAGGCCTGCC AGAGGGTAGA CATGGCTCAG
27551 GGGCTTGGAC AACATCAGGA AGTACCTCTA CATTTGCAAA TGCCTACTGT
27601 ATGCAAGGTG CCTCATTTCT CTGGACCCCC TTTTGCTTTT CTGTGAAAGG
27651 AGACAGACCA AATGATCCTT AAGGCTCCCT GACATTGTCA GTGATTGCCA
27701 AGGCAAACCT TGGCACGCTG CTCCCTGTTG AGAAGCAGCA TGGGGCCATG
27751 AGCTTTCAAG GCTGCTACAT CCAGCCTTGA CTGTTCTGCC ATTTAGGAGC
27801 TATGGGACCT TGAACAAACC ACATAACCTC TATGAGCCTC TGTGTTCCTC
27851 ATCTGTAAAG TGGGGGTGAT GACACCTTCC CTGCAAGGTA GATGTGAGGT
27901 CAAGAGGAAA CAAGGTACCT GGCCTACCAA GACTACCAAG AGCAGGGTCT
27951 TAGGAAATAG CTCTTATTCC ATCCTTGATG GGGCCTGTCC TTGATAGCTG
28001 GGCTTGGAGG CAAGGTGCTA ATGGGCAAGA CAAGAACTCC TATCGGGGGC
28051 TGGAAGTCAT TAAAGCTCTT GAACCCTGGT AGGAAGTTGC CATGTTCTGA
28101 GGGCACAGGG CCTCCCACAG TTTGAGTGAT TATTGCTATG AGAGAGGAGG
28151 TTCTCCAGGG AGCTGAGGAG TCCTACACCT GGGCTCAAAT GGATTTGCTG
28201 CAAAGGTGAC CAGCTGGTTC CCATTCTTGC AGGGAACCCT GCTCTCTCGC
28251 AGGGAACGGC TCCTCCAGAG TCTGTCTGTA TCGTGTTCCA TGTTGTCAGG
28301 GTTGCTTCCA GCCGGTTGGC CCTCCCCAGC TTTCCCACAG ACTTCCCACA
```

```
28351 CTGGAGCCCT GAGGGAGGGT CCTAAGCAGT TGCAGGAAGA GCTGAGAGGC
28401 CCCCGGAACT TGAGGAGCGA TTCCAAACCC AGGGACAGAG CCATCGTGGC
28451 TGGTTTCCTA AACTCCAGTC TCCTGTCTAC CCAGTCCTGC TCTGGAGAAA
28501 TCCCAGGGAC CACAGGCTTG GGAAGGAGGA AGGGGAATAG GCGTTCTGTC
28551 CACAGGGAGG TCCAGGCAAC AGCTTTCCCT CTTTCTCTAT GAACAATCAT
28601 CCTCTGGACC TCAGGGCTCC TGAGTTAGCA TTCTGTAACC TGGGTCCAAG
28651 AATCAGCCAA AGGTGTATTG TGGGGATACT TGTGTGTCAC CCCCCGCCCT
28701 AGGTAAGGCA GCACAGGCTG CAGGCCCCTG GGGTAGTGGC CTGCTCTGTG
28751 TGTCAGAGCC AGCCTCCCAG GAGGACAGAG CCACAGTGCC CCAGGCAGCC
28801 TCAATACAAC ACTCCCTGTC TCACAGGAAC AACTGCACTT GCCTCAGACA
28851 CTCGCAAAGC TCCCTGGAGA AGCAAGAACT GAAGAAAGTG AACTGGAACC
28901 TCCAGGTAAG CCCTGCAGCC CTTCTCTTAC TGACCCAGCT GGGGGGGCCCC
28951 CTGTACTCCA AGGACTGGGA AATCGAATGC CCAGCAGGAT GTGGCCAAGA
29001 GCAAGCCACT CCCTAAAAGC AGATTGCAGC CCCTGAAATA CTTACCCCTG
29051 CAAATTGAAC ACCAAGGCCA GGGAAGGGAG TGAGAGACCC CAAAGTGGAA
29101 GCTGAGAAAA TCCCCTTCTC CCAGCGGGTA GGCAGCAAGA GATTCCCAGA
29151 GTAGACTCCT TGTGGTAGGG CCCATTCCCC ACCCAGAGCC ATGTGTAATA
29201 ATTACTACTC ACTTCCTCCC CTCCCTTCAT TAAAAACAAA AGGCTTAGGC
29251 CCGACACAAT GGCTCACGTC TGGTGTCCCA GCTACTCAGG AGGCTGAGAT
29301 GGGAGGACAG CTTGAGCCCA GGAGTTGGAG GCTGCTGTTA GCTATGATGA
29351 TGCCATTGTA CTCTGCCTAG ACAACAGCGT GAGACCCTAT CTCAAAAAAA
29401 AAAAAAGAA AAAAAGAAAA AGGCTTAGCC CTGCCCTACT TAACTCTACC
29451 TCAAATTCTC CTTGCCCTCT CTCTGCCCCC TTCCATCTCC CCACCTCCAC
29501 TCCTGCTTAT GTCTCTGCCT CTATTGTTCC CTCTCAGGCT CAGGTAGCAT
29551 TTCCATTCTG CAAACTGACC CTCCTTCATT CACAAGGCAA GTCTGCTTCC
29601 CTCCTCTAAG GAGCTTCCCC TGCCTGAACT TCACCCGCGG ACATCTCCCC
29651 ATATCACATT CAGTCTGTAC TTGATGGGCC CTAAAAGCCC CAAAGGGTTC
29701 TCATGTTTTC ACATCTTGGC TCATTTTTCC AGATGGATGA TAAACTCCTT
29751 GAAGATAAGT ACATCTAGTC TGTTCCTTTT ACATTCCATG CTTGGGTACT
29801 TAAATCCAGC CACCGTGGAC TCTCCTCCCG CAAAGTTCAT GGGCATTTTG
29851 GGAGCTGGTG TTGAGATGCT CCCCATCTGA CCTGCAGCCC CATGTTCTAA
29901 TTGACCTCTT CGTGCAGTGA GAGGAGGGGA GGACTTTGGC CTATGCAATC
29951 TGGTCAGTGG CTCAGACCCA GCCTTTCAGG CAGAGGCTTT GGAATGGGAC
30001 TGGGTGGAGC TGTGTAGCTA GGGAGCTTCT CCCACCAGGA GCCGCTGGGT
30051 TCAACTCATC TCTGATCCTG AGAACCAGCA TAGGGCTTTG AAATGTCCGT
30101 GCCCATGAAT GGGTGGAGAA TAAAAGTATG TTTGCATCCC ACTAGAGTAG
30151 CCCCTTAAAG TCACTGTCCT TTAGGGTGAG TTGACTCCCG TCAACAACCA
30201 ATCCAAGGCA GCAGGACTGG ACCCTGTCTG TGCAGCCTTG CCAGGAGGGT
30251 TGAGCAGCTT CTCTCTCTGT CCCCAGCATG GCATCTCCAG TTTCTCCTAC
30301 TGGCACCAAT ACACACAGCG TGAGGACTTT GCGGTTGTGG TGCAGCCTTT
30351 CTTCCAAAAC ACACTCACCC CACTGAACGA GGTGAGCTGC AGGTATTTTA
30401 GGGAGGCTCA CGTATGGGGG CCTTATCACA GACGATGGAT GTATTTCCTT
30451 CTCTAAGTGG GCTTTTTTTT TTTTTAACC ATCTCTCTCC AAGAGGATTC
30501 CTGAGGGTGG CTTTTTCCAC ATTACCTCCT TTTTGTGGGG GCTGGGCTGT
30551 GATTGGAACT CAGATGTACT TTGAAAGGAA ATCAATAGTG ACTAAGCTCC
30601 CAGGCCTGGC CCTGATGTTT TCTGGATTGG GATAGAATGG AAAGCTTCCT
30651 AAAAATGTTA CTCTTTTCAA CTCTTAGGAT AGGGGTGCTG AAAGAAAAGG
30701 GAGAGACTAT GGGTGGGTCC AATTCTTGTC TGTTTAAAAA GAAAATTCCG
30751 GCCGGGTGCA GTGGCTCATG CCTGTAATCT CAGCCTTTGG GAAGCCAAGG
30801 CGGTTGAATC ACGAGGTTAG GAGTTTGAGA CCAGCCTGGC CAACATGGTG
30851 AAACCCCGTT TCTACTAAAA ATACAAAAAG TTAGCTGGGC GTGGTGGCAG
30901 GCACCTGTAA TCCCAGGTAC TCGGGAGGCA GAAGTTGCAG TGAGCTGAGA
30951 TTATGCCACT GCACTCCAGC CTGGCTGACA GTGCGAAACT CCGTCTCAAA
31001 AAAAAGAAA AAGAAAAAA AGAAATTCTA AATTCTGGGA GTTTTTCCAT
31051 CAGTATCTGA GCAAGTTGGC AGGAAAGTTG AAAGAATGAA AGGAGACATG
31101 CCCAGGGCAC CTGCTGGGAG AGTGAGTGGG GCTCAGGTAG CAGAGCCCTT
31151 TCCCAGGATG ATAACCTCCT TGCCGTTGGT TGCAGAGAGG GGACACTGAC
31201 CTCACCTTCT TCTCCGAGGA CTGTTTTCAC TTCTCAGACC GCGGGCATGC
31251 CGAGATGGCC ATCGCACTCT GGAACAACAT GGTGAGCAGC CAAGGGCCTG
31301 GTGGGCCTTG TCAAGGGGGG ATCTAAGGAT ATTGACACTC TGTCTCACAA
31351 TGGCAAAACT ACTGGAGACA TGGCTCCTTT CTCCCAAAGG CCCAAAGTGG
31401 CAGCACACCT TATTGGTCCT GATAGATTAA TTCCAAAGGG AAAATACCCT
31451 ATATTTATCC AACACCCTTT GAAAGTTATA CAAACACACA CTCACACAAC
```

FIGURE 3, page 10 of 33

```
31501 TTTATTCTTT GTTCCTTCAG CAATGCCCAG GTACTGCGAG GGGATCCCTT
31551 TGTAATCAGA TAGGTTGGCT AGATGAAAAT ACCAACTTCT ACCTCGTACT
31601 GTGTGACCTT GGGCAAACGA TCTCTCTGGC CACCTGTATC AACATCTATA
31651 AAACAGTGAA AACAAGACAG GTCTCAGACA ACGCATTGAG ATCATGTGTA
31701 CATGGCACCT AGCACAATAG TTAGCACTCA GCAAATGTCA CCACCATCAG
31751 CCTTCCAAGC ACTCCGGGCT CAACTCATAC CCAACTCATT TCTCTAAACA
31801 TCGAAAAGTG GAGATCCACA CAGCCTGTTT TCCGAGGCTG ATACCTATTC
31851 CAGTCCTTTC TGATGGGAAG AAGGGACCTT ATGAAATGAA CATACAGTCT
31901 GGGGGTCTTT CAGGGACACC TGCCTGGTGC TTCCACTCTG CCTTCTGTGG
31951 CTGGCCACCA GCAACTGAAC GGTTTCCGCA CAGCACTTGA CCTGTCACCC
32001 CCAACAACTG GATCCTCTTG CACGGAGCAA ATGAAATGCC TTCCCAACCC
32051 AATGGTTTCT TTTAATCCAG GCTCAGTGGG TAACACAATC CCCACCCCAA
32101 CCTGTATGTT CCCTCCTTTG TCCTATGACA ACTAAACAAG CTACATTCCA
32151 GCTCCTTTTA TCACAGTTTC AGGCCCGTAG TGTCTCTGCC AACCACCGCT
32201 GTGCAAACGT TCCCACCCCT GTCAGCTCAT CCAGTATGTC CAGCATCCCA
32251 CTCGGCTGAC TCACAATATT GACTTTCTCC TTAGCTATAC CATCTCCTCC
32301 TCTCTAGCAA CCTCTTCTTT TAAGAACAGC ATGTAAACTG GCTTTATCCT
32351 TGGCCTAGTT AATGGCAGAC TCAGCTTATG TCGACTTCCA TTGTCAGGGG
32401 GTTTTCCTCC TGTGGACATC ACGTACCTGC CCACTCCAAG AACTTCTATT
32451 GTACTCTTTC AGCCCAAGAC TCCGGATTGT AACCAAAATG TCTTTGAGTT
32501 TGCCCCACGA TTTTTAAATC AGTTTATATG GTATAATTCC TGTTTCTTTT
32551 GTGGATTTTG TTTTGAAGGC GGTTGTCCTT CACTGGCTGA ATCATGTGAC
32601 TTTATTCCTT TGTAAAAATC TTCCCAAAGA AAGGGTACCT ATTCCCTGTT
32651 CCTTTTCCCC TGAGACCTCA GGGGATTCCA CAGATGCCCT TGGCCCTTCC
32701 TTCCCAGTTT TTTCATCAAG GTATGGCCTT CCTACCAGGT GGCACTCCAA
32751 GTCTGCTTAA ATCTGGGACC CTCCAGGAAT CTCCTGGGGC TGGATAGCCA
32801 TAGTGACGGC TGGAACATGA AAAAGAGTCC ATTGGTTTCT TTTCTTGTGA
32851 ATTAACAATG TAGCTCTGGC CAGGCACGGT GGCTCATGCC TGTAATCCCA
32901 GCACTTTGGG AGGCCGAGGC AGGTGGATCG CTTGAGCCCA GGAATTAGAC
32951 ACCAACCTGG GCAACACAGG GGAGATTCTG TCTCTACAAA AATAATCAAA
33001 ATATTAGCCA GGTGTGGTGG TGCATGCCTG TAGTCCCAGC TGCTCAGAAG
33051 GCTGAGCGTGA GAAGATCACT TGAGCATGGG AGGTCAAGGC TGCAATGAGC
33101 CGAGATGGCA CCACCGCACT CCAGCCTGGG CAATAGAGTG AGACCCTATA
33151 TCTCAAAAAA CAAATAGAAA AAAAAAATAT ATGTAGCTCT GGCCTTCTCT
33201 TCTAAAGCAG TTCAGTAGCT CTTCCCATTC ACCCAGGTAA GAGGCCTTTA
33251 TTTCATAAAG ATAAGTGGGA GGAGTTTAGA TATGAAAACA AAACGTAAAC
33301 ACCGCACTGG AGCTATTGTG GAAACAAAAC AAGACTGTCC ATGGTTCCCC
33351 AGCCATTATT ATCTCAGCCA TACCCCGAAT TTCAAAATAA CAAAAACAAA
33401 ACTAAAGCCA TCCAGGGGTT TCTTATCCTA GGCTCTATAA TTTGGGTAAA
33451 TAATTATACA GTCTAATGTT TTCATCCAAA GCCAATCTTA GACATAAAGC
33501 TGTAGCATGA TGCCAACTTT TCAGATCGGC TTCTGGCTGG AATTTCACCC
33551 CTAGAGTAAC AAAAAATAAA TAATAGACCA TTAGAGCTGG AACAGACTGA
33601 GAGGTCATCT AGCCAGAACA TTCTGTAACT AAAGCATAGA AACATGAAGC
33651 AGTTTGCCCA ACATAACACA GACTGTTCAT GGCACAAGGG GGATTACAGA
33701 CCAGGTTTTT CTAGTCCTTT CCTGGTGACC TGGGCATGCC ACCACCCTCC
33751 CCACTGCTCC CAACCTGATA AGCACATATA TACCCGGTGA ATTCATGTCT
33801 CACAATTAGA GTCCTATGAC ATAGTGTCTG CAGGCTTTGG CTGATGTTCC
33851 CATAGTGTCT GCAGGCTTTG GCTGATGTTC CCAGGGTTCC CTACTAGGAA
33901 GCAAAAAGCA CCTTAAACTA TTTCATCTTA TTTCATCTCC TGCCCCTCCT
33951 CTCACGTCCT TCTCGAGACT TTTGCAAAGG CAAAGCCAGA AGCTCCAGCA
34001 GCACCAGGGG ATATTTTCCT CTTCCTCTGC CTTCTTCTGT CTTCTTATCT
34051 GAAGAAGTTT CTCTTTCCCG AGGCCTAGTC CTCTACTGCT GCCTCTACTC
34101 CCTCTTCTGC AGAAATCCTG CTCTCAGCCA GTGTTTGTAT CTCCCCAGGT
34151 GCTGGGTGAC AGCTCCAGCC TCCTAACTGA CATCCCTGTC TTCAGACTTA
34201 GAGCTCTTAG AATCGTGACT CTCAGCTCTG GCTGCATATT AGAATCATTC
34251 AGGGACATTG TGTATGTGTG TATGTATGTG TATATATGTA TGAATGTGTG
34301 TGTATGTGTG TGTGTATG TATGTATGTG TATGTGTGTA TGTATGTATG
34351 TATGTATGAC AGAGTCTCAC TCTGTTGCCC AGGTTGGAGA GCAATGGCAC
34401 CATCTCAGTT CACTGCAACC TCCGTCTCCT GGATTCAAGC GATTCTCCTG
34451 CCTCAGTCTC CCAAGTAGCT GGGGTTATAG GTGCATGCCA CCATGACCAG
34501 CTAATTTTTG TATTTTTAGT AGAGACAAGG TTTCGCCATG TTGGCCAGGC
34551 CAGGCTGGTC TTTAACTCCT GACCTCAGGA GATCCACCCA CCTCGGCCTC
34601 CCAAAGTGCT GGGATTATAG GTGTGAGCCA CAATGCTTGG CCATCCAGGG
```

FIGURE 3, page 11 of 33

```
34651 ACTTTTAAAA CAATTAGTGC CTACAGCCAC TTTGGAAAAT TCTTTGGTAT
34701 ATTTAATAAT GCTGAACACA TGTATTTCCT GTGATCCAAG GATTTCACTC
34751 CTAGGTATAT CCCCAAAAGA AAGGTATATA TGTGTCTACC AAAAGATACA
34801 CACAAAAATA TTCACAGCAG CACTATTTAT AAATAGCCCC AAACTAGAAA
34851 CTCCCCAAAT GTCCACTGAC AGTAGGATGG GTGAGTAAAC TGTGGCACAT
34901 TCATACTAAG GAATACCATA CAGGTCTGAC CGCATCTGTG GCTTTAAACA
34951 AAAATCAAGC AGGGTGATGT GACACAGAGT AATGGCTGGG AAGAGGGAGG
35001 CCTCACTGAA GAAGTGACAG CTGAACAAAC TTCAACAACA TACAATAATA
35051 TCTATAAAGT TCAAAAGCAA GCAGCTTGGC ATATGGGGTT AGACGTCAGC
35101 ATGATGGTGT AGAGACTCAC TGGGGGATGA ATAGTCCTGG AAGAAGGTGG
35151 AAAGGGGCTT TTGAGGACTA TAATAGTCTG TTGCCTGACT GGATGCTGGT
35201 ATGTTCATTT TATCGAAACT TATCTGTTGC TCACTTATGA TTTGTACTCG ,
35251 TTTCTATGTG TATGTTAGCT TCAATTAAAA GTTTACTTGA GGCCGGGTAC
35301 AGTGGCTCAC ACCTGTAATC CCAGCACTTT GGGAGGCCGA GGCAGGCAGA
35351 TCCCCTGAGG TCAGGAGTTC AATACCAGCC TAGCCAACAT GATGAAACCC
35401 CATCTCTACT AAAAATACAA AATTAGCCAA GCGTGGTGGC ACGTGCCTAT
35451 AATTCCAGCT ACTTGGGAGG CTGAGACAGG AAAATCGCTT GAAACCAGGA
35501 GGCAGGGGTT GCAGTGAGCC AAGATTGCAT CATTGCACTC CAGCCTGGGT
35551 GACAAGAGTA AAACTCTGTC TCAAATTTAA AAAAAAAAAA AAAAAAAGT
35601 TTACTTGAAA AACAATATCA GTGCCTGACC GGGCTTATCC CCAGAGAGTC
35651 TGACTTAATT GGTCTGGAGT GCGAGCTGGA TTCGGTACTT TGTGAAAGCT
35701 CCTGAGATTA TTTTAATGTG CAGGGTTTAT GAACCGCTGC CTTAGATCTG
35751 GTCCCCACAG AGAAATCAAG TAATCTGTAT AAAAGAAAAC CTGACCCAGT
35801 CACTCCCCTG CTTTCAAACT TCCAAAGCCT CCCACCTCTG AAGGAGGCAG
35851 GCCAGCCCCC ATAGCACAGC ACACTAGGCC TCTGGGACTT GGCCTGGTTC
35901 ACCTGATTAA CCTCTCTGGC TACCATTTCC ACCAGCGTCT GCCTCGCATG
35951 TTACAGTCTA GTGACTCCAG CAGCGTCCTG CACCACCTGT GGTGTTCCAC
36001 ACCTCTGCTA ACTCTTGCTC TCCTCCTTCT CCTGGATTGC CCTTCTCACC
36051 TCCTTGCCCA CTCCACCACT CAACTCAGGT GCCACCTCCT GCAGGAAGCT
36101 ACCTCTGAAT CTCCAGGACA GGCCAGTGGC CCACCCAGGT CCATTACACC
36151 CTGCCCAGTC CTGTCATTTG CTACGTGGTT GGTAGCCACA GTGCCTGGCT
36201 TAGGAAAGAC TGGTTCTAGG AAAAACAATT TCATTCCCTG TGGCCAGCTC
36251 CAAGCCTTCC CCCGCCAAGC TTCTCCATTC AGGTCTCTGT GAATTTAATT
36301 AATTCATCCA TCCATCAAAC AAGTATTTAC TGAGCACTAA TATGTGCTAG
36351 GTACTGCTCC AGGTGCTGAG GACTCAGCAG TGAAAGATG ACTGCTACTC
36401 TCATGGGACA TACAGGATAG TAGGGAAAAG ACAGATAATC AACAAGGTCA
36451 TTTCTGACCA CATCTGTGGT TTAAGAAAAA GTCAAGCAGA GTGATGTGAT
36501 ACAGAGTAAT GGTGGGGGAG AGGGAGGCCT CCCTGAAGAA GTGACAGTGA
36551 ATTGAGAAGC GCATGTCAAG GGGTTGCCAG GCAGAGGAAA TAGGACCCAC
36601 ATGGGCCTAG AGTCAGGAGT GAGCTTGAAG TGTCTGAGGA ACTTAAAGGC
36651 CAATGTGACC AGAGGGAAGT GAACAAGGTG AAAAAGTTGG GCAGGGGCCA
36701 GGTCCCTAGA TGCTTCTAAG CAGTAGAGTG ATATGCTCTG GCTTACCCCT
36751 GGGTCCGTGT ACCCTGGACT GGAAGAAAGC AAGGGTGGAC CTGGAAAGAC
36801 CACTAGGAGG CTGCTGTTGA TGGGTGAGAG AGGAAGGGGG CTGAGAGTAG
36851 GGTCAGGGCA GAGGAGGAGA GACGCTGTCG TGGGCTGGCG GATGGATGAT
36901 GGGGAAGAGG AACAAAGGAT GACTTTTTGG TTTGGGGTCT AAGAAACTGG
36951 GTGGATGATT GAGCAGGTAG AGAAAAAATC AGCGTGGGAG GAAAAAAAAT
37001 CAAGACTTCT GTTTTGGACA TGGTGCAAAC TGCCTTCCAG ACATCCACAT
37051 AGAGGTATCA GGATACAGAA GTTTGGAACT CACAGAGGAA GTCAAGGCTG
37101 GAGATTGAAA AAAAAAAAA AAAAAAAAA AGTGGGGTTA TTAGCATAGA
37151 GGGCCAATAT GGTGAAACCC TGTCTCTACT GAAAATACAA AAATTATCCA
37201 GGCATGGTGG CATGCACCTG TAATCCCAGC TACTCAGGGA GGCTGAGGCA
37251 GGAGAATTGC TTGAACCCAG AGATGGGGTG GAGGTTGCAG TAAGCTGAGA
37301 TCGTGCCATT GCACTCCAGC CTGGGTGACA GGGCAAGATT CCATCTAAAA
37351 AAAAAAAAG CCACTACAGG ATCAACTAAG AGCTCCTAGA GAAAGAATAG
37401 GTAGGTAGAA AAGAGTGTAA GGCCAACTAC CTAGCCCTGG GCATTCATTC
37451 CAGCTTTCAA CTCCAGTGAG AGATGAGAAG GAGAGTGTGG AGGTAGATGG
37501 GAAATGAGAA ACAATGCTGT GTCCAGAGAG CTAAGAGAAG TCAGTGTTTC
37551 AAGAGAGACA GAGCTGTCAA CTTTGATGGA TGCTTCTGAG AAGCCAAGCA
37601 AGTTGAAGAC AAAAAAAAAA AAAATGATCT TTGGCTCTGC CCATATGGCG
37651 ATCGTTGGTG GCCAGGGCCA GAGCTTCCAT CCAGCGATGG AGACTGCAGA
37701 CTGGCTGGAG CGAGCAGCAG AGAGAAGGAG AGATTAGGAA GTGCTGCCAG
37751 CACCTATAGA CAGCTCTTCC CAGAAGTTAT GAGAAGTAAC AGCCACGGTC
```

```
37801 ACTGGAGGGG ACATGGATCA AAGAAAGGGC AGGTGAAGGA GGGGAGATGT
37851 CGGAGCAGGT TGTGTACTGA CGAGAAGGAA CCAGTAGAAA GGGAGAAACT
37901 GATGCACTCA TCAAACCCTT GTAATCACGA TCATCTTCTG TGTGAATTAG
37951 TTCTGGGTTC CTGGAATAGC ATCGGGAATC AGCCGCGCTG ACCTTTAGCA
38001 TTTATTCTGT CACTGTTACG ATAGACTTGA GTTTCCTCAG TTCTTAAGAA
38051 AGTGGAAATA ATACTACCTT ATGTATGTAA GCCACCCCAA TCACACGTGC
38101 TTTCATGCCA TCTTCTCATT TGATGCTCAC AACAAACCAA GGTGTCAGGA
38151 CAGGATGTCA TGCCTGTCCC TAACCTCAGA GAAGTAGCCA GCCCAGCACT
38201 GCACAGCTCG TTACCAGCAG AGCCTGGATC CCAGCCATCT GCCCATCGTG
38251 CTCAGTCTAG TCACCCTAGC ATCTCTCCCA GGAACAGAAC TGCCCTCCCT
38301 CCTCCAATTG TGTTACTAAG GAACGGGTGT AAAAGGCCTT CAGACAACAG
38351 AAAGTGAGAT CATGGGCCAG GTGTGGTCAT GAGCTCAGAT AGTGAACTTT
38401 CACCTCCCTC CCTGGCAATA CCCTGTGGTC AGGAGCAGGC AGATTACACA
38451 AAGAGTGGAG GCTAGACGTT CCAAACAGAC TCTGAATAGG TGACAGTGCC
38501 AGGGGCTCAT TCTTCTCAGT GCTGGCCACA GGTTGGGCCT GGCTGCTGGC
38551 TAAAAGGTGC CGGGGAGGGG GATACAGCAG CTCCCAGCTC ATCCTCAGAG
38601 GGTCCTGGGA TCAAAGGTAT TTACACCCAG GGATATTTCA GATAAATCTT
38651 TTCATCTATG TGGAAAACAT ACAAAGTGGC GCAAGTGAGA AACTCCGATT
38701 TCCTAAGGTT GACAAGTCAA GTGCAGTAAT GATGTCATGG TAACCAATAT
38751 GTTTCCAAAC TTTCCTAAGG TTGACTAGCC CCATGCACTT TGAGAAGTTG
38801 GTAAATAGGA TTGTCGTCGT TTTATAAAAT TGAAAACACG GTGTCTTGCA
38851 ATCACAGCCA CTCACAAAGG AAGCCAGAGA TGGTCCCAGC CCCTCCGCAG
38901 ACTTCCTGTG GACTCAGGAC TGGTGGTCTC TCCTGGGCCT TGCTGTACCC
38951 GGCAAATCCA GGGGCACAGA CTCAGGGTTC TGCCCTGCCG ACAGATGCTG
39001 CCTAGCCTTC TGTGTGTCAT AAGTCAACTC CCGCTCAGCC CCAGGCTGCT
39051 GGGTCCCTGC TGTGGGCCAA AAACCAGCCA CTTCGCTGGT TTCTATCCCC
39101 CACCCCGTTC CCGAGGGAGG GGCTCTGGTG TGAGACACCC CCTCAGAGAG
39151 GAAAGTGTCT CCCAGCTTTG GAGAGAATCG AGGTGTCCTT TCTCTCTCTC
39201 CAGCTGGAAC CAGTGGGCCG CAAGACTACC TCCAACAACT TCACCCACAG
39251 CCGAGCCAAA CTCAAGTGCC CCTCTCCTGT GAGTAAACGT CCTGCCTGCC
39301 CCAGGTGGAA CAGATGCCTG GGGTGGGGGT TGTCCTGTCC CCTGGAAGCA
39351 CAGAGGAGTC CCCGGGGATG CTCCCTCAAA TGCGGCTTCA CTCACTGCCG
39401 TCTTCTCAAA TCCCACCTGT CCCCAGTGCC ACGGAAACTT CTCAGTGTGT
39451 GGGCAGCCAT GGAGGGAGGG GAGAGGACGT TCAACAGCTC CAACCGAAGG
39501 GAGGACAGTC GCTCAGGGAG GCAGATGAGC ACTGGCGGGT GTCTCGGGTC
39551 ACCCATTCCT TCCGAAAGCT CTGATGCATC CTCAGTCTTA AAAGTGCACC
39601 AAGGCCAGGC GTGGTGGCTC ACGCCTATAA TCCCAGCACT TTGGGAGGCC
39651 GAGGCGGTCA GATCACCTGA GGTCTGGAGT ACAAGACCAG CCTGGCCAAC
39701 ATAGTGAAAC CCCGTCTCTA GTAAAAATCC AAAAGTCAGC CGGGCATGGT
39751 GGCGGGCGCC TGTCATCCCA GCTACTCAGG AGGCTGAGGC AGGAGAATCA
39801 CTTGAACCCG GGAGGCAGAG GTTGCAGTGA GCTGAGGTCA TGCCACTGCA
39851 CTCCAGACTG AGTGACAGAG TGAGACTGTC TCAAAAATAA AGTGCATCAA
39901 GCAGCTGTCC CGTGCCAGGC AGTATACTAG GATCTGGGGA TCGGGAGGCA
39951 AAGATAAAAT AGACTCAGTG TCTGTTCCTG GAGCCTGCAA TGGTCTTCCT
40001 CCCTCGCCAC ACCCACTGCC CTTGCCTGGC CCACCTTCGA AGCCTGTGAC
40051 TTGTCTCCCC AGCTCTCCTC TCCCTCTTCT CCATCCACCC TACACTTGCT
40101 GCCAGACACA GATAGACCTT CCTGGAAATA ACTTGCCCCA TCAAGGCTGC
40151 TTGAAATCCT TGCCTGATCC CTACTGCCCA TTGACCAGAG TCTGGAGGGA
40201 GGGTCACCTC CCTCCATGAT ACACACTGCA CTCCTGGCCG GTGGATCCAT
40251 CTCCCAGGAA GCCCCACGAC TGCCCGCATC CAGGCCTTTC CTTTTGCCAT
40301 CTGTTCCTGG AGGTTCATCT TCCATCTGCT ATGAGAACAT CCGCCTCCCT
40351 CCAGGTCCAG ATGTTGCCTT TACTAAGCGA TGGTTTCACC GTCTCTTACC
40401 TACCATTCCT GTCTCCAGAC ACTGACCCAT GTGGGTCTCC TTTTCTATTT
40451 GTACCTCTCA TGAGACACCG ACCCAGTCTC CTTTATGATG TGATTGTTTC
40501 TGCACATCTC AACTTCCTCC TGGGCCACAA GAAAAGATGT CACATCTTAA
40551 CCCTCCAGTC TCATCACAGC TTCCAGCAAG GGGGCTAAAC ACAGCACGTG
40601 CCCAATTCAC ATTCACTGAG AGGAGAGTGG AGAGGGGCAT AGGAAGGCAA
40651 GAACGCACAC GATCTGCCCA CATGCCTCCC CTCCCGGCCC TTCTGATTTG
40701 GGGATCTTTC ATCTACTACA AAACCAGCTG TCCTTCCATG CTGCCCTTCC
40751 CTGATTTCTG GGTAGTCCTG GGATGGGAGA ATGGGGACAG TTGTGACCAC
40801 GAGGAAGCAG AGGTGGGAGT TCTACAGGCC CCACAGGGCT CTCTGCCATT
40851 GGTCACCTAT CAGTTCCCAA TCTTTCAAAA TCAGGTTTGA TGGCCAAGGA
40901 AACGCTGGTG AGAAACCAAA AGAAGGTTCT AGCTGGGTGT TGACCTCTTT
```

```
40951 AGAGGCCCAT CCCGCTAAAG AGGGTTTGGG CACAGCCTAA ATGAGGGAGC
41001 TTTACAAAAG GGAAGCTCTG TGAAAACGTG CAGGGTTATC GCAGCATCTC
41051 AGGAATGGGG ACTAGGCAAG TCTTGGCTTG GTGATGGATG GTTCACGGAG
41101 ATCCTTTCCA CTGACCCCCG CTCCTCCTCC ACAGGAGAGC CCTTACCTCT
41151 ACACCCTGCG GAACAGCCGA TTGCTCCCAG ACCAGGCTGA AGAAGCCCCC
41201 GAGGTGCTCT ACTGGGCTGT CCCAGTGGCA GCGGGAGTCG GCCTTGTGGT
41251 GGGCATCATC GGGACAGTGG TCTGGAGGTG CAGGAGAGGT GGCCGGAGGG
41301 AAGATCCTCC AATGAGCCTG CGCACTGTGG CCCTCTAGGC CCGGGGGTGG
41351 GTCCTCACCC TAAACTCCCT ATAGCCACTC TCTTCACCGC CCTCTGCCCC
41401 AGCCACTCCC GGCCACCAGG ACATGCTTCA ATGCCTGGTG CCATAGGAAG
41451 CCCAGGGGAC AGTCACAACT TCTTGGGGCC TGGGCTTCTT CCAGGCCTAT
41501 GCTCCTGGAA TGGATACATT TAAATAAAGT CCAAAGCTAT TTTATTCCTG
41551 GGTTTGCCTG CGTGAAGCAC TCACCTTCCA TCTCTTGTGC AGCCCAGGTG
41601 TGGGAGCTGC CACTTTTTGT GGCCTGCCTC CAGCAGGGCT GCCCAAGCCA
41651 CGACCAACCA GAGCCCAAAC TGCCTGCCAC CACGAGCATA TCCTCAAGTC
41701 ACCAAACCCA CTATTTCAAA GGCAGAAAAA ATGCTGGTCA CCAGGTGGTG
41751 GCTGGAATTT TGGAGCTGGC TGGTTGCCAT TCAGTCCAAT CCAACACATA
41801 CCTATTAAGC AACTGTTTTG TATCCAGGAC AATGCGAAGC ACTGAGGTGC
41851 CTCCTAGGCT GTGCATGTCG CAGCCTGGCA GAGAGGTCAA ACTCCTTCAA
41901 TAACCAAGAA GCCACGTGAT GATGTGTAAC TACTAGGGCA TCAGTAGGTA
41951 AATGTGTCTG ATTGTTTTAA AGAATAGAAA GGGTTCTTCG GGGAAAGTTT
42001 CTTGGGGGAG AGCAACCTTC ACATGTCATT TTGGGAAAAG GAATAAAAAA
42051 TGATTGGGAC ACAAATACCT CCTATATTCT CAACCTGATT TTCTCAAGGT
42101 GCTAAATTTA GGAAAAAATT CCTATTTCTA TATGCCCAGG TTTCTGAGGG
42151 AAAACTAGAG AGAGTCTGAA AATATGGGCT GCATTCACTG AGCCCCTGCT
42201 AGGGGCGAGG CCCCGTGCTG GAGGCCTTCC ACAGATGGTC TCTTTTATGC
42251 TGCACAAAAG CCCAGGGAGG GGGTAAAAGG AAAATCTTTG AAAATAGAAG
42301 TGATGCTTGC GCAACACCGT GAATGTACTA AACGCCGCGA ATTGTTCCAT
42351 TTAAAATGAT TAATTGTGTA TCATGTGAAT TTCACTTCAA TAAAAAAGAA
42401 TCCAGGGAGG TAGACATCAT CTGCATTGTA AACCTCTCTC TGATCCTGAA
42451 GTCCGGGATG ATAAAGAGCC TGAGTCACAA TCCCGGATGC AACACTGAAA
42501 TGCTGTGCCC TGAAGCTGCC TTCGCCAGCC TGAGCCCAGT GTCCCAGGCT
42551 CTGCATCTGT AAAAACTGGA GTAAGAGTAC ACATTTTGCT TATCTCACGG
42601 CGCTGCTGAA AAATAAGGAA CCGTGTGTGA ACCTCTAACT CTAAAATGCT
42651 GCACAACTGA AAATGGCCTT TTTCCTCGGT GAAGAGTTGG GATAAGGCCC
42701 AGACTGTTGG GGAAGATGTG AGACCCAGAG ATGAGTTTGG GGAAATGGGG
42751 TAATAACATA TGGGTGGAGA GTGCCCGCCT TCCTCTCAGG GAGGTTCATC
42801 ACCTTATCTC TTTCTGTCAC AACAGAGAAC CCGGAGGACC TATACCCAGT
42851 TCCGTGTTCT TCTGGGCTTC AGTGTCTGTT TCTATACAAT GGGAACAGCA
42901 TGCATTCCCC TGCTTTTTCC TATAGACTGG AAAACGTGGT GACCAAGTCA
42951 CACATCCCAG CTTATGCTCC CGGCTTAAGA CAGTGTAACG ACAAAGGTAA
43001 CCCTTACACT CCTGGTTTGA GACAGTATAA CGACAAAGGT AACATAGGAA
43051 GTCAAGGAGT TCGCTTCACC GCCCCTCCCC CCACCCCACC CTTTTTTTTT
43101 CCTGCAAGTT TCTATTCTTC CCGCAGCTCC TACCTCAAAG CAGCATGGAT
43151 TCATAACCAC AGGCTCCCCT CATTAGGGCT TGGGGAGGGA GGGTGTTGGA
43201 ATCCACACTG CCAGAGTAAT CCAGACTAAA ACATCAACAA ATGGTCCCAG
43251 CTGGTTCACC AAGGAACACT TGGCAAAACA AAGAAATCCT GTCTGGAGCG
43301 ACACGGACAC AGCCACAAAC CAGTCACCAA ATTCCAGCA AGTATGTGCT
43351 AAGAAGCCAA AAATTAAAAA TACGTGAGAA GCACCACTT GAAATTGGTG
43401 GTATTACATA CATACACTGG CTGTGCGCCA TGGGGTTTTT CTGTAGGAAA
43451 ATGTCCAGTC TAGCTAGAAC GGCACCCACA GCCACACCAT GAGCAAAGCC
43501 ACCAAATACC TGAGGAAGCC ACAGTCCATG GCACTCCCCA TGG
```

FEATURES:

```
Start:   3000
Exon:    3000-3098
Intron:  3099-4565
Exon:    4566-4637
Intron:  4638-12578
Exon:    12579-12644
Intron:  12645-16277
Exon:    16278-16338
Intron:  16339-16891
```

```
Exon:      16892-16950
Intron:    16951-18968
Exon:      18969-19036
Intron:    19037-19303
Exon:      19304-19347
Intron:    19348-19582
Exon:      19583-19682
Intron:    19683-23297
Exon:      23298-23349
Intron:    23350-24668
Exon:      24669-24749
Intron:    24750-26803
Exon:      26804-26869
Intron:    26870-27334
Exon:      27335-27426
Intron:    27427-28826
Exon:      28827-28905
Intron:    28906-30276
Exon:      30277-30381
Intron:    30382-31185
Exon:      31186-31281
Intron:    31282-39203
Exon:      39204-39278
Intron:    39279-41134
Exon:      41135-41335
Stop:      41336
```

CHROMOSOME MAP POSITION:
Chromosome 2

ALLELIC VARIANTS (SNPs):

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
1573	A	T	Beyond ORF(5')			
1737	T	G	Beyond ORF(5')			
2498	T	C	Beyond ORF(5')			
3097	C	T	Exon	33	A	V
3116	A	G	Intron			
4823	G	A	Intron			
4924	A	G	Intron			
4989	A	G	Intron			
5274	C	T	Intron			
5792	–	T	Intron			
5871	G	T	Intron			
6562	C	T	Intron			
9859	G	A	Intron			
9875	A	G	Intron			
10279	C	T	Intron			
10474	–	G	Intron			
10566	C	T	Intron			
12038	G	A	Intron			
12159	G	A	Intron			
12979	A	G	Intron			
13865	C	A	Intron			
13964	A	G	Intron			
14087	C	A	Intron			
14309	G	A	Intron			
16028	G	T	Intron			
16375	C	T	Intron			
16705	G	A	Intron			
19708	G	C	Intron			

21406	G	A	Intron			
22401	T	C	Intron			
22926	A	T	Intron			
23007	G	A	Intron			
23180	C	G	Intron			
26490	A	G	Intron			
26505	A	G	Intron			
29336	T	C	Intron			
29829	C	T	Intron			
29830	G	T	Intron			
29840	C	T	Intron			
29944	C	T	Intron			
30468	T	–	Intron			
30471	T	–	Intron			
30802	G	A	Intron			
30894	G	A	Intron			
30907	G	A	Intron			
31447	C	A	Intron			
31603	G	A	Intron			
31685	A	G	Intron			
31833	C	T	Intron			
31970	C	T	Intron			
32177	A	G	Intron			
33018	T	C	Intron			
33090	T	C	Intron			
33993	C	T	Intron			
34284	G	A	Intron			
34314	G	A	Intron			
35392	A	G	Intron			
35599	T	A G	Intron			
35997	C	A	Intron			
36085	C	T	Intron			
36270	C	T	Intron			
36481	G	A	Intron			
36619	G	A	Intron			
37088	G	C	Intron			
37204	G	A	Intron			
37485	G	A	Intron			
37624	–	A	Intron			
37685	C	T	Intron			
37769	C	T	Intron			
38897	G	A	Intron			
40155	A	G	Intron			
40355	G	C	Intron			
40486	T	C	Intron			
40512	A	C	Intron			
40622	–	A G	Intron			
40654	T	C	Intron			
40933	C	G	Intron			
41171	T	C	Exon	418	L	L
41379	T	C	Beyond ORF(3')			
41388	C	T	Beyond ORF(3')			
41880	A	C	Beyond ORF(3')			
42278	G	A	Beyond ORF(3')			
42339	G	A	Beyond ORF(3')			
42612	A	G	Beyond ORF(3')			
42817	T	G	Beyond ORF(3')			

FIGURE 3, page 16 of 33

Context:

DNA
Position
1573 TGGCTGTCCCCACACTGGAGATGCCCTCACCTCCTGGTCTGGCCCACATGCAGTGGTGAT
    GCCTCAGGGTCTTTGTGACTTGGTCTATCCATGTGTCCAAGTCTGTAAAGGAGGACTTCT
    GCCAGAACGTCCCCTTCCAGAGGCTGGAGCCATGACTCCCCTGTTACCCAACTTCAAGGT
    GCCTGGCAGGAACTTCTATGATACCAGGCAGCCACAGAGGGGAGGGATCAAAGTTGGGAC
    AGAGGCTGGTGTTTGAGAGACAGGATAGCCTAGACTGTGAACATGGGCAGTGGTTAGGGA
    [A,T]
    GTAGACATATGTGGTCAAACTGTAACAGAAAGCAAGGAAAAGGTACAAGCAACTCAGTTA
    CCTTTAGGGGAAGAAGAGAATTAGGAGGGACACAGGGAGCTTCAAACTGGGAGTGTTTTG
    TTTCTTAAACTGGGCCATAAGTACATGGATGTGTGTTTTATTATTCTTTATATCTTACAC
    ATCTATTTACTCAGCAAATCTTACAGAACTTCCTGTGTACCAGGCATTGTTTCAAGTGCT
    TTAGAAATCTCTCTCTTAAGTAGATGTGATGGGTGTGAAATAATTCATGATGAAACCAAA

1737 TACCCAACTTCAAGGTGCCTGGCAGGAACTTCTATGATACCAGGCAGCCACAGAGGGGAG
    GGATCAAAGTTGGGACAGAGGCTGGTGTTTGAGAGACAGGATAGCCTAGACTGTGAACAT
    GGGCAGTGGTTAGGGATGTAGACATATGTGGTCAAACTGTAACAGAAAGCAAGGAAAAGG
    TACAAGCAACTCAGTTACCTTTAGGGGAAGAAGAGAATTAGGAGGGACACAGGGAGCTTC
    AAACTGGGAGTGTTTTGTTTCTTAAACTGGGCCATAAGTACATGGATGTGTGTTTTATTA
    [T,G]
    TCTTTATATCTTACACATCTATTTACTCAGCAAATCTTACAGAACTTCCTGTGTACCAGG
    CATTGTTTCAAGTGCTTTAGAAATCTCTCTCTTAAGTAGATGTGATGGGTGTGAAATAAT
    TCATGATGAAACCAAAGGGGACACAGTAGGGCACTCATGTGAAAGAAGGAGAGGTCTAAG
    GCATAGCATCAGAGGCCCCAAAAATATCAGCTCCAACACCAGAGGATGCATTTTCTTTTTA
    ATTAAACACTAAATTTTCACTGCCCAAATTCATTTGCTCAGCTGAATAATCGGTTGCAGG

2498 AGGGCGGGAGGGAGCCTCTCCACCCCTCCCCTGAACCTGGGCAATCAGAACCAGCCCCTG
    ATGGAAGCCTGAGCTCTGGGGCCTCCTGCCTCCCCCTCTTTGTGCAGCGTTTTGTGTAAC
    TGCGTTCTGACCCTGCGGGAGAACTCCCAAGAGCTAGCCAGGCTGGAGGCCTTCAGCCGA
    GCCTACCGGGTAAGACCAAGAAGGGCACCATGCTGTGTCCTCTCCCCTACGTTCACTCTA
    ACACACAGCCCAGAGCCCCTAGAGGAGGCACACAGGGAAGGAAAAGCTGGTCAGGGATTG
    [T,C]
    GGGGAGACGGGGAGCAGCCTGGGTGCCTTCCTCTGTCTCACGTGACTGTGGTGTCTCAGG
    TGCCCTGGTTGGAATCATCCCAGTAGGATCCAGGTGGAAAAGCCCTCATGGCCCAGCTAC
    CGTTGAGGGCTTAACCCCAACTCCTGGCCCGTAGCCCTGGATGCCTCATGAGACCACCTT
    TCCCTCCCCCACTCCCACTCCAAAGGCAGGTGCCGAGCCTCTGGAGGTTCTTCCCAGGTT
    TTTATCCCTTTTGGGACTTCCTGCCTAGCCCTTCAGAGAGAGTAGTCTACTTACAATCAA

3097 AAAACAAAAAGGTGACCCAACCTGTTTCCAAATTCTCTGGAAAGGGACTTGCCCTCAGGT
    GATTTGTGTTCTCAAGGGAAAGGCTGAGTCGGCCCCTCCATCCAGGGAGATGGACTGCCC
    ACCACCCCTACTCTTGCCTCACTGGGTCCTGGGCCCACCCAGGGCCTGGGCTGAAGACCC
    TGTGCATGTGTCCCCAGAGCAGCATGCGCGAGCTGGTGGGGTCAGGCCGCTATGACACGC
    AGGAGGACTTCTCTGTGGTGCTGCAGCCCTTCTTCCAGAACATCCAGCTCCCTGTCCTGG
    [C,T]
    GGTATGTCCCCTGCCCTCGCCCATGGTACTCTTTTAGAGGAAGAAATGCAAGGCAGAATT
    GCCAGTTGCTTCCACGAGCATGTGCATAAAATGGGAAAGACACAGCTCTCCAGACGCTG

3116 ACCTGTTTCCAAATTCTCTGGAAAGGGACTTGCCCTCAGGTGATTTGTGTTCTCAAGGGA
    AAGGCTGAGTCGGCCCCTCCATCCAGGGAGATGGACTGCCCACCACCCCTACTCTTGCCT
    CACTGGGTCCTGGGCCCACCCAGGGCCTGGGCTGAAGACCCTGTGCATGTGTCCCCAGAG
    CAGCATGCGCGAGCTGGTGGGGTCAGGCCGCTATGACACGCAGGAGGACTTCTCTGTGGT
    GCTGCAGCCCTTCTTCCAGAACATCCAGCTCCCTGTCCTGGTGGTATGTCCCCTGCCCTC
    [A,G]
    CCCATGGTACTCTTTTAGAGGAAGAAATGCAAGGCAGAATTGCCAGTTGCTTCCACGAGC
    ATGTGCATAAAATGGGAAAGACACAGCTCTCCAGACGCTG

4823 AACAAACTACTACACCCGTGTCTCTCTTTTCTTCCCTGATCAGCTTGAACCACTTGGAAG
    CAAAACAGAGACCCTGGACCTGAGAGCAGAGATGCCCATCACCTGTCCCACTCAGGTAGT
    AGGGGAGGACCTGCCTGGCTCCTCTCCACAAACCAGGGCACACAGCTCGCCCTACCCACT
    TCGTCCTCCACCACAGCTTCCTCAGTACCCATCTTGCCCCCTTACTGAGGCCTGAGAGAT
    TTGGAGGATGGAGGGGAGTCCATGAGGATGGACAGGGGAGGTGAGAGGGGGAGACAAGAGT

[G,A]
CAGCTGTCATTGGGAACAGGAGATGCAGCAGGGAGAGGAGGCCTGGGCCCCAGCAGAGGG
AGAGGATCCCGGTGAGAAAAGTGGGCTCCTGAGAGAGGAAATCAGGATGCCAGGAAAATG
GCAGGAGGGCTTCTCTTAGCAGTGGTGTTTGGGGCAGATGAAAAAATCTGACTGCAGGTT
AGAGGGCCCAGGCAGGAGCCAGGCAGGCTTAAGAGCTGTGGTTGGAGAGAGGAGAGCCTG
GATTAGGGAGATTCCACAAGGAAAGGATCACAGAGGACAGCAGCAAAGGGCAGAGCCCAG

4924   CCTGTCCCACTCAGGTAGTAGGGGAGGACCTGCCTGGCTCCTCTCCACAAACCAGGGCAC
        ACAGCTCGCCCTACCCACTTCGTCCTCCACCACAGCTTCCTCAGTACCCATCTTGCCCCC
        TTACTGAGGCCTGAGAGATTTGGAGGATGGAGGGGAGTCCATGAGGATGGACAGGGGAGG
        TGAGAGGGGAGACAAGAGTGCAGCTGTCATTGGGAACAGGAGATGCAGCAGGGAGAGGAG
        GCCTGGGCCCCAGCAGAGGGAGAGGATCCCGGTGAGAAAAGTGGGCTCCTGAGAGAGGAA
        [A,G]
        TCAGGATGCCAGGAAAATGGCAGGAGGGCTTCTCTTAGCAGTGGTGTTTGGGGCAGATGA
        AAAAATCTGACTGCAGGTTAGAGGGCCCAGGCAGGAGCCAGGCAGGCTTAAGAGCTGTGG
        TTGGAGAGAGGAGAGCCTGGATTAGGGAGATTCCACAAGGAAAGGATCACAGAGGACAGC
        AGCAAAGGGCAGAGCCCAGAGCTGTATGGAGGAGGGACGAGGGTGGGCCTACCAGGACAC
        GGCAGCTCCAGGCTCCTTTTAAGGAGGAATCCGTAAGTGGTTGTTAAGCTTGACTTCAGG

4989   TCGCCCTACCCACTTCGTCCTCCACCACAGCTTCCTCAGTACCCATCTTGCCCCCTTACT
        GAGGCCTGAGAGATTTGGAGGATGGAGGGGAGTCCATGAGGATGGACAGGGGAGGTGAGA
        GGGGAGACAAGAGTGCAGCTGTCATTGGGAACAGGAGATGCAGCAGGGAGAGGAGGCCTG
        GGCCCCAGCAGAGGGAGAGGATCCCGGTGAGAAAAGTGGGCTCCTGAGAGAGGAAATCAG
        GATGCCAGGAAAATGGCAGGAGGGCTTCTCTTAGCAGTGGTGTTTGGGGCAGATGAAAAA
        [A,G]
        TCTGACTGCAGGTTAGAGGGCCCAGGCAGGAGCCAGGCAGGCTTAAGAGCTGTGGTTGGA
        GAGAGGAGAGCCTGGATTAGGGAGATTCCACAAGGAAAGGATCACAGAGGACAGCAGCAA
        AGGGCAGAGCCCAGAGCTGTATGGAGGAGGGACGAGGGTGGGCCTACCAGGACACGGCAG
        CTCCAGGCTCCTTTTAAGGAGGAATCCGTAAGTGGTTGTTAAGCTTGACTTCAGGCCTGG
        GGTGGGGGCAGGTTCTCATTGTCTTCAGCTCCTGTTTCTAGGCCCGGTCTTATGGCTTTT

5274   GGGGCAGATGAAAAAATCTGACTGCAGGTTAGAGGGCCCAGGCAGGAGCCAGGCAGGCTT
        AAGAGCTGTGGTTGGAGAGAGGAGAGCCTGGATTAGGGAGATTCCACAAGGAAAGGATCA
        CAGAGGACAGCAGCAAAGGGCAGAGCCCAGAGCTGTATGGAGGAGGGACGAGGGTGGGCC
        TACCAGGACACGGCAGCTCCAGGCTCCTTTTAAGGAGGAATCCGTAAGTGGTTGTTAAGC
        TTGACTTCAGGCCTGGGGTGGGGGCAGGTTCTCATTGTCTTCAGCTCCTGTTTCTAGGCC
        [C,T]
        GGTCTTATGGCTTTTTAACCAAATAAGGCCAAGGCCAGAAAACCCTCAGCAGCAATAAAA
        GCAGAAGGCCTGACCCAATCTGGGAGGCTGGGTTTCCCTCCTAGGTCGGCCACACCACCC
        TCTCCCACCCTCCCTGCTGGGGAATGGACCTGCAGCTCCCCCATGTGTCTGCTGGGAATC
        CTGAGAGAGTGGGCACCCCTGTTCACATGCCTGCTCCCTGTCTGCTGCCTGCCCTACCCC
        AGTCTTGGGCTCAGGCTCAGTCTTGTGTGCCATCAGCCCCATCAGGAGAGCAAGAATGGC

5792   CTGTCTGCTGCCTGCCCTACCCCAGTCTTGGGCTCAGGCTCAGTCTTGTGTGCCATCAGC
        CCCATCAGGAGAGCAAGAATGGCAGGAAGAAGGGATGGGAAGTGAAGACAGTCGTAGCAG
        AGGGCTCAGTTGCTGGGTCTTGTGCTTGGAGCTAAGGAGATTGTCAGATTCTGCAACAGC
        TAGTGCAACACAGATGCCTCTAGTCCAGGTGGTCAGGTGCTGGCCAAAGGCCTGGAGCAA
        AACCTTAGAGGCCCCTACTGTGCCAGGTGTAAACTCTTTAACTGCTTTCCTAAGGATGCC
        [-,T]
        TGGGGGTTCTAGGGGAGCAGCCAGGGACCGTGGATAGTGGGGGCATTTGGGGACTCAGAA
        ATAGCCATATTGTAGATATTTCAATATTTTACCAACCCTATAGCCATACTGAATATCAGC
        CATGGAGGGCCCTTTCCAAACTGTCCACTCCCCTTCCATTACATAACAAAAGCAGCCATC
        ATTTGCTCTTTCTTTCAACAAACGTGTATTGAGTACTGAGTTGGAGCCTAAGCACTGGGT
        CAGGGAGAGCCCTGTCACCCTGGGCTTCGAGGCAACCACTTCCAGGCTTTACCCCAGATC

5871   TGGCAGGAAGAAGGGATGGGAAGTGAAGACAGTCGTAGCAGAGGGCTCAGTTGCTGGGTC
        TTGTGCTTGGAGCTAAGGAGATTGTCAGATTCTGCAACAGCTAGTGCAACACAGATGCCT
        CTAGTCCAGGTGGTCAGGTGCTGGCCAAAGGCCTGGAGCAAAACCTTAGAGGCCCCTACT
        GTGCCAGGTGTAAACTCTTTAACTGCTTTCCTAAGGATGCCTTGGGGGTTCTAGGGGAGC
        AGCCAGGGACCGTGGATAGTGGGGGCATTTGGGGACTCAGAAATAGCCATATTGTAGATA
        [G,T]
        TTCAATATTTTACCAACCCTATAGCCATACTGAATATCAGCCATGGAGGGCCCTTTCCAA
        ACTGTCCACTCCCCTTCCATTACATAACAAAAGCAGCCATCATTTGCTCTTTCTTTCAAC

FIGURE 3, page 18 of 33

```
         AAACGTGTATTGAGTACTGAGTTGGAGCCTAAGCACTGGGTCAGGGAGAGCCCTGTCACC
         CTGGGCTTCGAGGCAACCACTTCCAGGCTTTACCCCAGATCAGGCAGAGACCCCCAAAAG
         GAGGCTGCTCCACCCAGCAGCATCTTAAGCTGAGTGGGCTCAGTGCCTCCCTTCTAGACA

 6562    CTAAGCAACATTGGAGCCCATTCTGAAAGGGTCCATCTGTTGGCCAGCCCAACTTCACTG
         TGTTCTGAGCATTCTGCATTCCTCAGTCCCATCTGCTCCCTCCCATGTGCCTTGGAGTGA
         TATAAAAGTCCACCAGCATCTCAGTGTGAGCTGACAGGGGCCAGGCAGCACCTATTTTTG
         TCCTAGATGTGTCTAAACATAGAGGCAACAGGCAACAGGCAAGACGCAGTGGGGGGCGGG
         AGGCAGGAGGCCGAGATGGCTGTGAGCATGAGCTTTCTCAGCCTCCTCCCTTCTCCCATC
         [C,T]
         GCAGTCTAACTGCTCATACGTTCTGTGTGCCAGGTAGGGTGACTTAACAGCACGCCATGG
         ATTTCTGTTGTAGTTTCAAGTTGGACAAATTCTTTTACAGACAACTTTTGACTAGCCTTC
         TGTGGACTGAGCCTATACTCTGCCTTAATGGGCTCTCTGCCCACTCCTTTCCTAACCCCA
         GGGCAGCTGGCTGAACACCTGGTCCTTTTCTTAGGTTTCATTCTTTTTGACCTCTCTGAA
         GCCCTTGTCAAAAGTCACCACCTCCCCCTTGAAATTCACTCCTTCCTGGGTTTGTGGACA

 9859    CAATAAACCCAGCTAAAAACAAGCCCAATAAAACCCAATAAAACCCATTAGACAGGAACA
         TAGGAGTTGGAAAAAAAAGAAAAGAAGGGGAGGGGGAAGAAAGCCCTGAGGCACCCCGGC
         TGCCTGTCTGCCACAACCCTGGGCTGTAATTGTTCTTGCCATGGCCTCAGTCTGCAACAC
         ATTCTAGTGTCTCCTTGACCTCTAGCCCTCTAGCTCTGCCTCCCTTTCCCCAACCTGTAG
         ATCTTGTGATCAAATAGATTCAATGAAACACATTGTCCAGTTGC
         [G,A]
         CTCGCAGCACTTCCAAAAAGGTCAAGTTTGTCCTTCCCTCAGTGCCTCCCATTCTGGTCA
         CGGTAGGACTGACTCCAGCCCCTGGACCCTAAGCTGAGTCTGGGCTCCTTTGACGTGCAG
         GGAGAATGCCACTGAGTCTTGTCTCTGAGGACCCTACCTCTCCAAATCTTGCCTCAGTTC
         CTCAGCAGGTACTACACTGACTGGCCATGCCATTCTCTGATGCTTCACTGCCTCAGCTTC
         TCAAGTCTGTCTCCCCACCTGAGCCAATTGTGAGTTTCTCTCTC

 9875    CAATAAACCCAGCTAAAAACAAGCCCAATAAAACCCAATAAAACCCATTAGACAGGAACA
         TAGGAGTTGGAAAAAAAAGAAAAGAAGGGGAGGGGGAAGAAAGCCCTGAGGCACCCCGGC
         TGCCTGTCTGCCACAACCCTGGGCTGTAATTGTTCTTGCCATGGCCTCAGTCTGCAACAC
         ATTCTAGTGTCTCCTTGACCTCTAGCCCTCTAGCTCTGCCTCCCTTTCCCCAACCTGTAG
         ATCTTGTGATCAAATAGATTCAATGAAACACATTGTCCAGTTGCACTCGCAGCACTTCCA
         [A,G]
         AAAGGTCAAGTTTGTCCTTCCCTCAGTGCCTCCCATTCTGGTCACGGTAGGACTGACTCC
         AGCCCCTGGACCCTAAGCTGAGTCTGGGCTCCTTTGACGTGCAGGGAGAATGCCACTGAG
         TCTTGTCTCTGAGGACCCTACCTCTCCAAATCTTGCCTCAGTTCCTCAGCAGGTACTACA
         CTGACTGGCCATGCCATTCTCTGATGCTTCACTGCCTCAGCTTCTCAAGTCTGTCTCCCC
         ACCTGAGCCAATTGTGAGTTTCTCTCTCTCCTCCTCTCATCCTGGCACCTAGAAATGCTC

 10279   GGGAGAATGCCACTGAGTCTTGTCTCTGAGGACCCTACCTCTCCAAATCTTGCCTCAGTT
         CCTCAGCAGGTACTACACTGACTGGCCATGCCATTCTCTGATGCTTCACTGCCTCAGCTT
         CTCAAGTCTGTCTCCCCACCTGAGCCAATTGTGAGTTTCTCTCTCTCCTCCTCTCATCCT
         GGCACCTAGAAATGCTCTCTAACGCTTGAGCTGCTCAACCAGCATGGGTCACTTGTTTAT
         AGCATGCTCCCAGATCGCCCTCTTTGTTGGTGAATGCTCAGGGAATGCTTACTGTTAACC
         [C,T]
         GAGACAAGCCCAAGTAGCTACATGGACCTGCCACCATAAGCCCTCTCCTGTCTTATGCTG
         TTGTAGAGGGTCCAGGGCTCACTTCTCCCACTTGGCCCTGAGTACCTCTCCTTGAAAGGA
         TGTCAGGGGCTGGGCGCAGTGGCTCACGTCTGTAACCCCAGCACTTTGGGAGGCTGAGGC
         GGGCGGATCACCAGGTCAGGAGATCGAGACCATCCTGGCTAACATGGTGAACCCCCCGTC
         TCTACTAAAAATACAAAAAATAAAATAGCCATTTGTGGTGGCAGGTGCCTGTAGTCCCA

 10474   TCTCTAACGCTTGAGCTGCTCAACCAGCATGGGTCACTTGTTTATAGCATGCTCCCAGAT
         CGCCCTCTTTGTTGGTGAATGCTCAGGGAATGCTTACTGTTAACCCGAGACAAGCCCAAG
         TAGCTACATGGACCTGCCACCATAAGCCCTCTCCTGTCTTATGCTGTTGTAGAGGGTCCA
         GGGCTCACTTCTCCCACTTGGCCCTGAGTACCTCTCCTTGAAAGGATGTCAGGGGCTGGG
         CGCAGTGGCTCACGTCTGTAACCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACCAG
         [-,G]
         TCAGGAGATCGAGACCATCCTGGCTAACATGGTGAACCCCCCGTCTCTACTAAAAATACA
         AAAAATAAAATAGCCATTTGTGGTGGCAGGTGCCTGTAGTCCCAGCTACTCGGGAGGCT
         GAGGCAGGAAAATGGCATGAACCCAGAAGGCAGAGCTTGCAGTGAGCCGAGATCGCGCCA
         CTGCACTCCAGCCTGGGCAACAGAGCAAGACTCCGTCTCAAAAAAGCAAGCAAGAAAGAA
         AGGATATCGGTTACCTGTTTCAGACAGGAATGCTGAGACCAGGGAAAGGGGAGACTTGTC
```

FIGURE 3, page 19 of 33

10566   CTTACTGTTAACCCGAGACAAGCCCAAGTAGCTACATGGACCTGCCACCATAAGCCCTCT
CCTGTCTTATGCTGTTGTAGAGGGTCCAGGGCTCACTTCTCCCACTTGGCCCTGAGTACC
TCTCCTTGAAAGGATGTCAGGGGCTGGGCGCAGTGGCTCACGTCTGTAACCCCAGCACTT
TGGGAGGCTGAGGCGGGCGGATCACCAGGTCAGGAGATCGAGACCATCCTGGCTAACATG
GTGAACCCCCCGTCTCTACTAAAAATACAAAAAATAAAAATAGCCATTTGTGGTGGCAGG
[C,T]
GCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAAAATGGCATGAACCCAGAAGGCA
GAGCTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT
CCGTCTCAAAAAAAGCAAGCAAGAAAGAAAGGATATCGGTTACCTGTTTCAGACAGGAATG
CTGAGACCAGGGAAAGGGGAGACTTGTCGGGTGCCTCAGGGAACCAGTATCTGAGCTGGG
GGCTGAGAGCTCTGTGTGGGTGGACTCTGTCCTCCCAGTCGCTGCTGAGTCCCTCTCTTC

12038   CTGGGAATACAGGCATGGGCCACTGTGCCAGCCTGGTTTTTTCTTCTTGTTCCCATTTTA
TTCTCACATTTTCAGACCATGGGCTTACTACTCCACTGAGCACATTTTGTGAGAGTGCTC
ACAGCCCTGGGCCCGGTTGCTGTTTCCTGATCTCAGTCTTATCAACTTGATCTTGCTTTG
CTGTCATTTATACATTTTCTCATTAGCTTTCTCCCCATTTCTTCTTTGTCTGCTTCCTTC
TTCCTTCTTTAACTAACTCCTCACCTGCAACTGGGGGGACTTGGATTCTTGACTGGGCTT
[G,A]
TGTGAAAACTGATTGTAAAACAGATAGGTAAGTAGGGAATGAGGAGGGTGTTTTACAAGA
AAAAAAAAATGACTAAGATACAGGAACCCAACCTAAAGAGGAAAAGACATACAGTTCAAA
GGAGGCAGAAAGAAAAACATTACAGATACTCAAATATATTGATAATCATAACACTTTCTG
GAAGATTAAAAAAAATGCTGAAACATGAATCCCTTGCTAGAGAAATTACAAAGCCAAGAAA
ATAGATAGGTCTGAGGATTAGGGAGCTGTTCAGTTGCTAGGAGGAACACAAAAGCACAGA

12159   CAGCCCTGGGCCCGGTTGCTGTTTCCTGATCTCAGTCTTATCAACTTGATCTTGCTTTGC
TGTCATTTATACATTTTCTCATTAGCTTTCTCCCCATTTCTTCTTTGTCTGCTTCCTTCT
TCCTTCTTTAACTAACTCCTCACCTGCAACTGGGGGGACTTGGATTCTTGACTGGGCTTG
TGTGAAAACTGATTGTAAAACAGATAGGTAAGTAGGGAATGAGGAGGGTGTTTTACAAGA
AAAAAAAAATGACTAAGATACAGGAACCCAACCTAAAGAGGAAAAGACATACAGTTCAAA
[G,A]
GAGGCAGAAAGAAAAACATTACAGATACTCAAATATATTGATAATCATAACACTTTCTGG
AAGATTAAAAAAAATGCTGAAACATGAATCCCTTGCTAGAGAAATTACAAAGCCAAGAAAA
TAGATAGGTCTGAGGATTAGGGAGCTGTTCAGTTGCTAGGAGGAACACAAAAGCACAGAC
CCCAGACTACAATGGGTATGAAACCCTCTGCACGCCTTTTGTTGTCCATCCCTTGCCAAA
GCTGTTATGTAAAACCCTCCGGGGGAATGAATGAAATTATGTTTATACAGTTCTTTCTAT

12979   CAAACCATTTCCACCTGCCAGGGGCTCGGGTGTGGTACAGGTTTCAGAGTATTCACTGAA
GCAGAAATGTACTTCTTACATACTGGGGATTGGAATGTACAGAAAAGGCTCCCGGACCAC
GAAGCCCCAGGATTGTCCTAACATGTTCTCAAGTTGCTTACCTGACGTCAGCCCCCAAGC
AGAGGAAGTGTCTATGGATCGATTTTCTTTGACCTTGGCAATCCTGGGCTCACAGACGTG
GTTACTGCTTAGGCAGCTCAGCCTCTCAAGAGGGAGAGGCAGCTGGTGTGATGTGGCGTT
[A,G]
ACTTCTTGGAAGGTGGAGGCTGAGTGGGAGGGAACTACAATTCTGGGGATGGGACCCAAA
AGGAAGTGGAGGCACGTTGTTCATGTTCCTGTGGGCCCCTAGGCCTTGTTTGGTTCAAGT
CAATCATTCTAGTGCTGAGGATTCAGAGCCCATGGTTAATTCCATTGGATTAACCATGTC
TGTGAGCCTAGGACGGCCACTGCAAAGACGGCCTGGAGGACCCCGGACTATACCATGACT
GGCAGTCAGGCCTGGTCCGGATCAGGTCTGTTGGTCACCAGGATGGGGTTTGACCCGCAG

13865   GGGCTGGGGCAAGGGAGGCAAATTCAGGTGCACCATGCAGGCCAGGCCTTCCTGAGGTGA
GATTTAAACTGAGACATGCATAATGAGGAGACACTTGCTATACAGGGAGCCAGGAACACA
GTCCCAGGCAGAAGGACCATGGACCACACAGGCTCAGAAGTGGGACTGTGTTGGGTGTAT
TTGGGGAAGAGAAAGAAGGTCAGAGTGGCTGGGGGCATGAGAATGAGGTGGAGAGTGGGG
GAAATGAGATCAGGAGTGCCAAGGAGCCAGATCACACAAAGCCTGAATTACTGAGTAAAA
[C,A]
CACTGGATTTCAAGTGGAGAAAGATGGGAAGGCATTGGCGGTCTCAGGAGAGAGTGACAT
GATCTGGTTCACGTCTTTCAAAGATCTCCCTGACTGCTATGTGTAGAATGGGTTGGCCAT
CAGCAGGAGTGATTGGGGAAAGACATTTTATAAGCCAGCTGAAGAAACTAACCCATATGA
AATCATTAAGAACTATTGGATGCTAAGCTCTGGGGTGCAAGCAATACCAGATTGCTGGCT
GCGGGTTATGCTGTGTCCAGCCTCTCTGAATTTTCTCAGGCTCACGTTAGCCCAGTGGAG

13964   ATACAGGGAGCCAGGAACACAGTCCCAGGCAGAAGGACCATGGACCACACAGGCTCAGAA
GTGGGACTGTGTTGGGTGTATTTGGGGAAGAGAAAGAAGGTCAGAGTGGCTGGGGGCATG

```
        AGAATGAGGTGGAGAGTGGGGGAAATGAGATCAGGAGTGCCAAGGAGCCAGATCACACAA
        AGCCTGAATTACTGAGTAAAACCACTGGATTTCAAGTGGAGAAAGATGGGAAGGCATTGG
        CGGTCTCAGGAGAGAGTGACATGATCTGGTTCACGTCTTTCAAAGATCTCCCTGACTGCT
        [A,G]
        TGTGTAGAATGGGTTGGCCATCAGCAGGAGTGATTGGGGAAAGACATTTTATAAGCCAGC
        TGAAGAAACTAACCCATATGAAATCATTAAGAACTATTGGATGCTAAGCTCTGGGGTGCA
        AGCAATACCAGATTGCTGGCTGCGGGTTATGCTGTGTCCAGCCTCTCTGAATTTTCTCAG
        GCTCACGTTAGCCCAGTGGAGGCTTGTCCTCATTGAACCAGTGACCAAATTCCCTGAGAA
        TTGAAACGTCAGCTGCATCTTGTGAATCAGGCATTTCTTCATTTATTCATTTACCTATTG

14087   ATGAGGTGGAGAGTGGGGGAAATGAGATCAGGAGTGCCAAGGAGCCAGATCACACAAAGC
        CTGAATTACTGAGTAAAACCACTGGATTTCAAGTGGAGAAAGATGGGAAGGCATTGGCGG
        TCTCAGGAGAGAGTGACATGATCTGGTTCACGTCTTTCAAAGATCTCCCTGACTGCTATG
        TGTAGAATGGGTTGGCCATCAGCAGGAGTGATTGGGGAAAGACATTTTATAAGCCAGCTG
        AAGAAACTAACCCATATGAAATCATTAAGAACTATTGGATGCTAAGCTCTGGGGTGCAAG
        [C,A]
        AATACCAGATTGCTGGCTGCGGGTTATGCTGTGTCCAGCCTCTCTGAATTTTCTCAGGCT
        CACGTTAGCCCAGTGGAGGCTTGTCCTCATTGAACCAGTGACCAAATTCCCTGAGAATTG
        AAACGTCAGCTGCATCTTGTGAATCAGGCATTTCTTCATTTATTCATTTACCTATTGGAT
        GCCTATGTAGAGTGGGCACTGCACTAAGTGCTCGGTAGACAGTGGTGAGCCGAATGGGTC
        TGGATCTGCCCTCTTGGTTCTTCAGTCTCATGCATCTTTGCTTTTGCTGCTGGAAGAGCT

14309   CATTTTATAAGCCAGCTGAAGAAACTAACCCATATGAAATCATTAAGAACTATTGGATGC
        TAAGCTCTGGGGTGCAAGCAATACCAGATTGCTGGCTGCGGGTTATGCTGTGTCCAGCCT
        CTCTGAATTTTCTCAGGCTCACGTTAGCCCAGTGGAGGCTTGTCCTCATTGAACCAGTGA
        CCAAATTCCCTGAGAATTGAAACGTCAGCTGCATCTTGTGAATCAGGCATTTCTTCATTT
        ATTCATTTACCTATTGGATGCCTATGTAGAGTGGGCACTGCACTAAGTGCTCGGTAGACA
        [G,A]
        TGGTGAGCCGAATGGGTCTGGATCTGCCCTCTTGGTTCTTCAGTCTCATGCATCTTTGCT
        TTTGCTGCTGGAAGAGCTAAAAATCCCAGAGCTAGAAGGGCGTGTGTTTGTTTTAACAGC
        TTTCTACTCAAAGTAACCACAGAAACAAAATTCTGTCATCTGAGGTAACGTGAATGAGCC
        TAGAGGACATTACGTTAAGTGAAATAAGTCAGGCACAGAAAGACAAATACTACATGTTCT
        CACCATATGCGGAAGCTTAAGAAGTTGACTTCACAGAAGTAGAGTATAAATAGTGGTTAT

16028   ATATGCAAAAAGACATACAGGTATCCAGAAAAGACAGGCAGAAACCAGGAGCTTTACAAT
        TTTAAAATATTTTGTGTTATTATTCTAAAAATATTTTAATTATTGTCTAGGTTCTACCAT
        TATAATTAGTGTCAGTTAGCTTAATTTTATAAAACACACATACCTGTAATCTCATGTTAG
        GCATCCAAATGCTGTGTTCCTTTGGGAGACCCACCTGTGTAGGACTTCATGGTTTTCTTC
        CCTGCTTTGGGGCAGCCACTGGCTCCATTCAAAGCATAGATATATGGGGATAAGAAAGGT
        [G,T]
        GTGTGTGGGTGCACATGTGGAGACATGCACTATGGGTTGTGCATAGGGGTAGCTAGACAC
        ACCCATTTCTCCCCCTTTAATTTCCCTCCTAGCCCACCTATAACTCACAGTTCTTTCCCT
        CACATGATCCTGTATGGTGACTCATTTCTAGCCTCCATCAAAAATCCCTTAGCTGGTTCT
        TCTTGGGCTGAAGCTTATCTCCCTGCACAATGAGTGTTGGGCACTGAATCTTTTCTCCTG
        TTGATTTAGAACTGGGGCAGTGACTTCCTGTGTACAGAGTGGAAGGCTTCCAATAGTGTT

16375   GGGTAGCTAGACACACCCATTTCTCCCCCTTTAATTTCCCTCCTAGCCCACCTATAACTC
        ACAGTTCTTTCCCTCACATGATCCTGTATGGTGACTCATTTCTAGCCTCCATCAAAAATC
        CCTTAGCTGGTTCTTCTTGGGCTGAAGCTTATCTCCCTGCACAATGAGTGTTGGGCACTG
        AATCTTTTCTCCTGTTGATTTAGAACTGGGGCAGTGACTTCCTGTGTACAGAGTGGAAGG
        CTTCCAATAGTGTTCCAACCTCTGGTGAGTGAAAACATCATCATCTCCTTCAATTAAGGG
        [C,T]
        CTTGCCGAATATCAGGTTGTGGGGAGACCCTGCAAACATACCCTGGAGCTTTAAGCAGGA
        CTTGCTAATTCCCCTGCAGTGCAGACCTAGATCCTGCGGCCTGCCGCCACAGCTGGGCTT
        CCATGTGGAGGTGCACAGAGCTCTCCATTGGATGCTACTTCTTGTCTCCTTATAGTCCCA
        GTGGCAGTCCCTTAGGCCTCCCTGCCCAGTGAGGCAGGTAGAGTCAGGGATTGGGATCTA
        CCTGCCTGTGCTACATGACCCTGCAGCTGGAACTTTCCTGGACCACCCCAATGTCAATCA

16705   CTGCAAACATACCCTGGAGCTTTAAGCAGGACTTGCTAATTCCCCTGCAGTGCAGACCTA
        GATCCTGCGGCCTGCCGCCACAGCTGGGCTTCCATGTGGAGGTGCACAGAGCTCTCCATT
        GGATGCTACTTCTTGTCTCCTTATAGTCCCAGTGGCAGTCCCTTAGGCCTCCCTGCCCAG
        TGAGGCAGGTAGAGTCAGGGATTGGGATCTACCTGCCTGTGCTACATGACCCTGCAGCTG
        GAACTTTCCTGGACCACCCCAATGTCAATCAGGCTCTTCTGAGGGTGGATGATAGCCATG
```

FIGURE 3, page 21 of 33

```
          [G,A]
          AACCCATTCCCTGCAGTGCCTTGGTTGGTCTGAATGAATGGGAGGGGCAAAACTGCTAAA
          GCCTTAAGCTGAAAATAAGTACAATGGGGAGCAGTGGGACAGAGTTATAGACTTCTGGTA
          AAATGTGTACTTTAAGAGGTAGATACCCCCAGCCCCCACAACCACCTCTCTGCTTGTCTC
          CCCTAGTCCACCAGCTCCGACCAGCAGACATCAAAGTGGTGGCCGCCCTGGGTGACTCTC
          TGACTGTGAGTAGTGAGCCATGAACCAGGATGGGCAGCTCAGAGTCCAGCCAGGCCCTGC

19708     CCTGCTGGAGGAGGGGAAGAGGAGGTTATCTGCAAGAAGGGAAGTCAGCCAGCCCTGAAA
          AGCCCCAGACTTCCTGTGTCCCACCCATGTCCCCACCCTGCATGCTCATCTCAGTTACTG
          TGAGGGTCCTGCAGGCTCTCACCTGTGCTCTTCTCCTCCTCCTCCTCCTCTAAAGACATT
          CTGAAGAAGTTCAACCCTTACCTCCTTGGCTTCTCTACCAGCACCTGGGAGGGGACAGCA
          GGACTAAATGTGGCAGCGGAAGGGGCCAGAGCTAGGTGAGTAGATGCCGTACAGGAGGGC
          [G,C]
          AG

21406     AAGAATGAATCTCCATCTCCAAAAATAATAATAATTAAAAATAAATAAAAGATACAAAGG
          AATCAAAAGATGAACTCCCTGGCCACGAAGAGCTTGCACTCTAGGTAAGGAGGCTAAACA
          AATGGGAAATAACTTTCTGAAAAAGACAATGCTGGGTATGGCAACAATGCAGTGCTTCGC
          ATGGAGTACAATTAAGAGAACAGAAGAGCACACAGTATGAACTGCACTGTCTAAAGACAG
          ATGCAGACCCAGAAGGGACCCCTGAAATCATCCAGTCCAACCTCTTCCTTTAAAAGATGG
          [G,A]
          AAAAGTCAATCCTAGCAAGATTCAGCAACTTGTACAAGCTCAACAGCAAGTTGGTAGCAG
          AGCTGAAAGTAGAACCACTGGTCCCTGGGGTAAAAAAGGAAATGCAAGATGTGTGGATCA
          GGGAGCCCAGAGAGGAGGCTCAAGGGAAAGTAGGACTTGGTCTGGGCCTGAAGGATGGGA
          AGAAGATGGCTAGGAAGAGGGGAAGAAGCGGCATTTGTAACTTCCCCTCCTACCCACGAG
          GGCTTATTGCCCATGGATTCTCTTAGTCACACCTTGAACCTGTTAAAAGGTTAAAGGCAC

22401     AGGCCAAGGTAGGCAGATTGCCCAGGAGTTGGAGACCAGTCTGAGCAACTTGGTGAAACC
          CTGTCTCTACCAAAAATACAAAAAATTAGCCGGGCATGGTGGCATGGACCTGTAGTCCCA
          GCTACTCGGAAGGCTGAGATGGGAGGATCACTTGAGCTCAGAGGGGTTGAGGCTGCAGTG
          AGCCGTGATCACACCACTACACTCCAGCCTGGGCAACAGAGCAAGACCCTGTCTCAAAAA
          GGATACAATTTAACATTGTACCTGTGAAATCATCACCACAATCAAGATGAAAAATGTGTT
          [T,C]
          ATCACCCACAGGAGTTTTTCTCAGGCCCCTTGGTAATCTCTCCCTCCTGCTCCTTCCTGTC
          CCTACCTCACACCCCAGGCAACCACTAACCTTCTTTCCATCACAATAGATTAGTTTGCAT
          TTTTAAAAATTTTATATAAATGGGATCAAAGAGTATATACTTTTTATCTGACTTATTTAG
          CAAAATGATTTTGCGATGCATCCATGTTATTCGGTATACCAATAGTTCGTCCCTTTTTAT
          GGCTGAGTGTAGTGTTCCGTTGGCATTCATATCGCTCATCCAGAACACCAAATGGTATTG

22926     GTTCGTCCCTTTTTATGGCTGAGTGTAGTGTTCCGTTGGCATTCATATCGCTCATCCAGA
          ACACCAAATGGTATTGTTTTATTTATGGCAGACATCAGGGGATGAAGGGAGAACTAATCC
          TGTCCATCCTGGTTTATTGGAGAGGGAGAAAAAAAAAAGTGAGGAGATGGGGAATGGTGC
          GGAAATCTAAGTAACCACAGAAAAGAAAAACAAAAGGATTAAAGGAGCAGAGAGCAGGGC
          TTAGAAGTAAAGGTTAAAGGAGTCATTAAGCCTGGAAAGGAGAAAACTGAGGGATAATTG
          [A,T]
          GAGCTGTGACTTTTCTCAAATATACAAAAGGTTATTTTTAAAACAGGCAACTGAAGAAGA
          AATGAACAGGCTTGGCTTACGAAGAAAGAGCTTGAGGAAGTATAAGGGAAAGTCCCTGAG
          GGGAGGCTTGACGGGATCCCAACCCGAGTGGCCGATGAGACTATTGGGTGGCAGGGGCTA
          GATCAATGTGGCTCCAGGGTCCAGGGCAGCCATGTGATTGTTACTAAGCTGAGATTTCTT
          GAGAATGGAATGACCTTTGTACTGGTAACATCATTCTTCTTGAAACACCTCTCTTCCTAG

23007     TTTATGGCAGACATCAGGGGATGAAGGGAGAACTAATCCTGTCCATCCTGGTTTATTGGA
          GAGGGAGAAAAAAAAAAGTGAGGAGATGGGGAATGGTGCGGAAATCTAAGTAACCACAGA
          AAAGAAAAACAAAAGGATTAAAGGAGCAGAGAGCAGGGCTTAGAAGTAAAGGTTAAAGGA
          GTCATTAAGCCTGGAAAGGAGAAAACTGAGGGATAATTGTGAGCTGTGACTTTTCTCAAA
          TATACAAAAGGTTATTTTTAAAACAGGCAACTGAAGAAGAAATGAACAGGCTTGGCTTAC
          [G,A]
          AAGAAAGAGCTTGAGGAAGTATAAGGGAAAGTCCCTGAGGGGAGGCTTGACGGGATCCCA
          ACCCGAGTGGCCGATGAGACTATTGGGTGGCAGGGGCTAGATCAATGTGGCTCCAGGGTC
          CAGGGCAGCCATGTGATTGTTACTAAGCTGAGATTTCTTGAGAATGGAATGACCTTTGTA
          CTGGTAACATCATTCTTCTTGAAACACCTCTCTTCCTAGGCCAAAATCCCATGTCGTGAG
          TCCTCGCTCCTGAGCCGGCACTAACGCCCCTCTCTCTACCCCCCACCTAGGGACATGCCA
```

FIGURE 3, page 22 of 33

```
23180    TAAAGGAGTCATTAAGCCTGGAAAGGAGAAAACTGAGGGATAATTGTGAGCTGTGACTTT
         TCTCAAATATACAAAAGGTTATTTTTAAAACAGGCAACTGAAGAAGAAATGAACAGGCTT
         GGCTTACGAAGAAAGAGCTTGAGGAAGTATAAGGGAAAGTCCCTGAGGGGGAGGCTTGACG
         GGATCCCAACCCGAGTGGCCGATGAGACTATTGGGTGGCAGGGGCTAGATCAATGTGGCT
         CCAGGGTCCAGGGCAGCCATGTGATTGTTACTAAGCTGAGATTTCTTGAGAATGGAATGA
         [C,G]
         CTTTGTACTGGTAACATCATTCTTCTTGAAACACCTCTCTTCCTAGGCCAAAATCCCATG
         TCGTGAGTCCTCGCTCCTGAGCCGGCACTAACGCCCCTCTCTCTACCCCCCACCTAGGGA
         CATGCCAGCCCAGGCCTGGGACCTGGTAGAGCGAATGAAAAACAGCCCCGTGAGTACAGG
         CCCCCAGGCCACCCCTGAAAGGTGCCCATCTCCTGCTGGCTGGGGAGGGGACAGCCCCAT
         AAGGGTCCCTCTCACCACAGCACTTCCTGCTTTGGGCTAGCCAAAAGATCCTCGGAGAAG

26490    CTCCCTCTGCTATAAAGCAAAGCCCTGAGATTCAGCCTGCAAGGACTTACTGAGCACCTA
         CTATGTACCTTGTTTGCATCACCCAGGATGCTGTGGACACACCTCTAAATCAGCCTCCTA
         CTGGGGAGATGGTTCAGAGGAAGAGAACCTTACACTGAGTCACAGGGGATAGAAGTTAGG
         GGAACACAGGAGAGCAAAACATTTCAGGCAGTGGGACCAGCATGGACCAAAGCCCAAAGG
         AAAAAGGAAGTGTGGCCACCCAGGGCATGGCAAGGGGCTGGAGAAGGCTGAGGTCAGATG
         [A,G]
         CGGATGGGACTGCCAAGAGCCAAGGCCAAAAAGTGGCAGGACCCAGCACTGGCAGAGTCC
         ACTGTTGGGTCTGAGATTATGTAGAGCAGGGTGGGGGTTGGGATTGTTCATGGTGTCTAG
         TAGGGGACAAGGGATGATTCCTTACAGAGACTCAGCAGCAACAAGAACTGGGCTTCTCAG
         TTTGACCAGGACCACCGAAGCCCCTCTGTACCCACTCAGTCATTTAGCCCAGGCCCCAGA
         GCCCTCCTATGCTCTTGCCATTCTCTCAGAGCGGGCACCAGGGGCTAAAGAGAGTACCCT

26505    AGCAAAGCCCTGAGATTCAGCCTGCAAGGACTTACTGAGCACCTACTATGTACCTTGTTT
         GCATCACCCAGGATGCTGTGGACACACCTCTAAATCAGCCTCCTACTGGGGAGATGGTTC
         AGAGGAAGAGAACCTTACACTGAGTCACAGGGGATAGAAGTTAGGGGGAACACAGGAGAGC
         AAAACATTTCAGGCAGTGGGACCAGCATGGACCAAAGCCCAAAGGAAAAAGGAAGTGTGG
         CCACCCAGGGCATGGCAAGGGGCTGGAGAAGGCTGAGGTCAGATGACGGATGGGACTGCC
         [A,G]
         AGAGCCAAGGCCAAAAAGTGGCAGGACCCAGCACTGGCAGAGTCCACTGTTGGGTCTGAG
         ATTATGTAGAGCAGGGTGGGGGTTGGGATTGTTCATGGTGTCTAGTAGGGGACAAGGGAT
         GATTCCTTACAGAGACTCAGCAGCAACAAGAACTGGGCTTCTCAGTTTGACCAGGACCAC
         CGAAGCCCCTCTGTACCCACTCAGTCATTTAGCCCAGGCCCCAGAGCCCTCCTATGCTCT
         TGCCATTCTCTCAGAGCGGGCACCAGGGGCTAAAGAGAGTACCCTTTTTTTCCTTACAGGA

29336    AAATACTTACCCCTGCAAATTGAACACCAAGGCCAGGGAAGGGAGTGAGAGACCCCAAAG
         TGGAAGCTGAGAAAATCCCCTTCTCCCAGCGGGTAGGCAGCAAGAGATTCCCAGAGTAGA
         CTCCTTGTGGTAGGGCCCATTCCCCACCCAGAGCCATGTGTAATAATTACTACTCACTTC
         CTCCCCTCCCTTCATTAAAAACAAAAGGCTTAGGCCCGACACAATGGCTCACGTCTGGTG
         TCCCAGCTACTCAGGAGGCTGAGATGGGAGGACAGCTTGAGCCCAGGAGTTGGAGGCTGC
         [T,C]
         GTTAGCTATGATGATGCCATTGTACTCTGCCTAGACAACAGCGTGAGACCCTATCTCAAA
         AAAAAAAAAAGAAAAAAAGAAAAAGGCTTAGCCCTGCCCTACTTAACTCTACCTCAAAT
         TCTCCTTGCCCTCTCTCTGCCCCCTTCCATCTCCCCACCTCCACTCCTGCTTATGTCTCT
         GCCTCTATTGTTCCCTCTCAGGCTCAGGTAGCATTTCCATTCTGCAAACTGACCCTCCTT
         CATTCACAAGGCAAGTCTGCTTCCCTCCTCTAAGGAGCTTCCCCTGCCTGAACTTCACCC

29829    CCCTCTCAGGCTCAGGTAGCATTTCCATTCTGCAAACTGACCCTCCTTCATTCACAAGGC
         AAGTCTGCTTCCCTCCTCTAAGGAGCTTCCCCTGCCTGAACTTCACCCGCGGACATCTCC
         CCATATCACATTCAGTCTGTACTTGATGGGCCCTAAAAGCCCCAAAGGGTTCTCATGTTT
         TCACATCTTGGCTCATTTTTCCAGATGGATGATAAACTCCTTGAAGATAAGTACATCTAG
         TCTGTTCCTTTTACATTCCATGCTTGGGTACTTAAATCCAGCCACCGTGGACTCTCCTCC
         [C,T]
         GCAAAGTTCATGGGCATTTTGGGAGCTGGTGTTGAGATGCTCCCCATCTGACCTGCAGCC
         CCATGTTCTAATTGACCTCTTCGTGCAGTGAGAGGAGGGGAGGACTTTGGCCTATGCAAT
         CTGGTCAGTGGCTCAGACCCAGCCTTTCAGGCAGAGGCTTTGGAATGGGACTGGGTGGAG
         CTGTGTAGCTAGGGAGCTTCTCCCACCAGGAGCCGCTGGGTTCAACTCATCTCTGATCCT
         GAGAACCAGCATAGGGCTTTGAAATGTCCGTGCCCATGAATGGGTGGAGAATAAAAGTAT

29830    CCTCTCAGGCTCAGGTAGCATTTCCATTCTGCAAACTGACCCTCCTTCATTCACAAGGCA
         AGTCTGCTTCCCTCCTCTAAGGAGCTTCCCCTGCCTGAACTTCACCCGCGGACATCTCCC
         CATATCACATTCAGTCTGTACTTGATGGGCCCTAAAAGCCCCAAAGGGTTCTCATGTTTT
```

CACATCTTGGCTCATTTTTCCAGATGGATGATAAACTCCTTGAAGATAAGTACATCTAGT
CTGTTCCTTTTACATTCCATGCTTGGGTACTTAAATCCAGCCACCGTGGACTCTCCTCCC
[G,T]
CAAAGTTCATGGGCATTTTGGGAGCTGGTGTTGAGATGCTCCCCATCTGACCTGCAGCCC
CATGTTCTAATTGACCTCTTCGTGCAGTGAGAGGAGGGGAGGACTTTGGCCTATGCAATC
TGGTCAGTGGCTCAGACCCAGCCTTTCAGGCAGAGGCTTTGGAATGGGACTGGGTGGAGC
TGTGTAGCTAGGGAGCTTCTCCCACCAGGAGCCGCTGGGTTCAACTCATCTCTGATCCTG
AGAACCAGCATAGGGCTTTGAAATGTCCGTGCCCATGAATGGGTGGAGAATAAAAGTATG

29840   TCAGGTAGCATTTCCATTCTGCAAACTGACCCTCCTTCATTCACAAGGCAAGTCTGCTTC
CCTCCTCTAAGGAGCTTCCCCTGCCTGAACTTCACCCGCGGACATCTCCCCATATCACAT
TCAGTCTGTACTTGATGGGCCCTAAAAGCCCCAAAGGGTTCTCATGTTTTCACATCTTGG
CTCATTTTTCCAGATGGATGATAAACTCCTTGAAGATAAGTACATCTAGTCTGTTCCTTT
TACATTCCATGCTTGGGTACTTAAATCCAGCCACCGTGGACTCTCCTCCCGCAAAGTTCA
[C,T]
GGGCATTTTGGGAGCTGGTGTTGAGATGCTCCCCATCTGACCTGCAGCCCCATGTTCTAA
TTGACCTCTTCGTGCAGTGAGAGGAGGGGAGGACTTTGGCCTATGCAATCTGGTCAGTGG
CTCAGACCCAGCCTTTCAGGCAGAGGCTTTGGAATGGGACTGGGTGGAGCTGTGTAGCTA
GGGAGCTTCTCCCACCAGGAGCCGCTGGGTTCAACTCATCTCTGATCCTGAGAACCAGCA
TAGGGCTTTGAAATGTCCGTGCCCATGAATGGGTGGAGAATAAAAGTATGTTTGCATCCC

29944   TCTCCCCATATCACATTCAGTCTGTACTTGATGGGCCCTAAAAGCCCCAAAGGGTTCTCA
TGTTTTCACATCTTGGCTCATTTTTCCAGATGGATGATAAACTCCTTGAAGATAAGTACA
TCTAGTCTGTTCCTTTTACATTCCATGCTTGGGTACTTAAATCCAGCCACCGTGGACTCT
CCTCCCGCAAAGTTCATGGGCATTTTGGGAGCTGGTGTTGAGATGCTCCCCATCTGACCT
GCAGCCCCATGTTCTAATTGACCTCTTCGTGCAGTGAGAGGAGGGGAGGACTTTGGCCTA
[C,T]
GCAATCTGGTCAGTGGCTCAGACCCAGCCTTTCAGGCAGAGGCTTTGGAATGGGACTGGG
TGGAGCTGTGTAGCTAGGGAGCTTCTCCCACCAGGAGCCGCTGGGTTCAACTCATCTCTG
ATCCTGAGAACCAGCATAGGGCTTTGAAATGTCCGTGCCCATGAATGGGTGGAGAATAAA
AGTATGTTTGCATCCCACTAGAGTAGCCCCTTAAAGTCACTGTCCTTTAGGGTGAGTTGA
CTCCCGTCAACAACCAATCCAAGGCAGCAGGACTGGACCCTGTCTGTGCAGCCTTGCCAG

30468   CCTTTAGGGTGAGTTGACTCCCGTCAACAACCAATCCAAGGCAGCAGGACTGGACCCTGT
CTGTGCAGCCTTGCCAGGAGGGTTGAGCAGCTTCTCTCTCTGTCCCCAGCATGGCATCTC
CAGTTTCTCCTACTGGCACCAATACACACAGCGTGAGGACTTTGCGGTTGTGGTGCAGCC
TTTCTTCCAAAACACACTCACCCCACTGAACGAGGTGAGCTGCAGGTATTTTAGGGAGGC
TCACGTATGGGGGCCTTATCACAGACGATGGATGTATTTCCTTCTCTAAGTGGGCTTTTT
[T,-]
TTTTTTTTTAACCATCTCTCTCCAAGAGGATTCCTGAGGGTGGCTTTTTCCACATTACCTC
CTTTTTGTGGGGGCTGGGCTGTGATTGGAACTCAGATGTACTTTGAAAGGAAATCAATAG
TGACTAAGCTCCCAGGCCTGGCCCTGATGTTTTCTGGATTGGGATAGAATGGAAAGCTTC
CTAAAAATGTTACTCTTTTCAACTCTTAGGATAGGGGTGCTGAAAGAAAAGGGAGAGACT
ATGGGTGGGTCCAATTCTTGTCTGTTTAAAAAGAAAATTCCGGCCGGGTGCAGTGGCTCA

30471   TTAGGGTGAGTTGACTCCCGTCAACAACCAATCCAAGGCAGCAGGACTGGACCCTGTCTG
TGCAGCCTTGCCAGGAGGGTTGAGCAGCTTCTCTCTCTGTCCCCAGCATGGCATCTCCAG
TTTCTCCTACTGGCACCAATACACACAGCGTGAGGACTTTGCGGTTGTGGTGCAGCCTTT
CTTCCAAAACACACTCACCCCACTGAACGAGGTGAGCTGCAGGTATTTTAGGGAGGCTCA
CGTATGGGGGCCTTATCACAGACGATGGATGTATTTCCTTCTCTAAGTGGGCTTTTTTTT
[T,-]
TTTTTAACCATCTCTCTCCAAGAGGATTCCTGAGGGTGGCTTTTTCCACATTACCTCCTT
TTTGTGGGGGCTGGGCTGTGATTGGAACTCAGATGTACTTTGAAAGGAAATCAATAGTGA
CTAAGCTCCCAGGCCTGGCCCTGATGTTTTCTGGATTGGGATAGAATGGAAAGCTTCCTA
AAAATGTTACTCTTTTCAACTCTTAGGATAGGGGTGCTGAAAGAAAAGGGAGAGACTATG
GGTGGGTCCAATTCTTGTCTGTTTAAAAAGAAAATTCCGGCCGGGTGCAGTGGCTCATGC

30802   TGAGGGTGGCTTTTTCCACATTACCTCCTTTTTGTGGGGGCTGGGCTGTGATTGGAACTC
AGATGTACTTTGAAAGGAAATCAATAGTGACTAAGCTCCCAGGCCTGGCCCTGATGTTTT
CTGGATTGGGATAGAATGGAAAGCTTCCTAAAAATGTTACTCTTTTCAACTCTTAGGATA
GGGGTGCTGAAAGAAAAGGGAGAGACTATGGGTGGGTCCAATTCTTGTCTGTTTAAAAAG
AAAATTCCGGCCGGGTGCAGTGGCTCATGCCTGTAATCTCAGCCTTTGGGAAGCCAAGGC
[G,A]

FIGURE 3, page 24 of 33

```
              GTTGAATCACGAGGTTAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCCCGTTTC
              TACTAAAAATACAAAAAGTTAGCTGGGCGTGGTGGCAGGCACCTGTAATCCCAGGTACTC
              GGGAGGCAGAAGTTGCAGTGAGCTGAGATTATGCCACTGCACTCCAGCCTGGCTGACAGT
              GCGAAACTCCGTCTCAAAAAAAAAGAAAAAAGAAAAAAAGAAATTCTAAATTCTGGGAGT
              TTTTCCATCAGTATCTGAGCAAGTTGGCAGGAAAGTTGAAAGAATGAAAGGAGACATGCC

30894         AAGCTCCCAGGCCTGGCCCTGATGTTTTCTGGATTGGGATAGAATGGAAAGCTTCCTAAA
              AATGTTACTCTTTTCAACTCTTAGGATAGGGGTGCTGAAAGAAAAGGGAGAGACTATGGG
              TGGGTCCAATTCTTGTCTGTTTAAAAAGAAAATTCCGGCCGGGTGCAGTGGCTCATGCCT
              GTAATCTCAGCCTTTGGGAAGCCAAGGCGGTTGAATCACGAGGTTAGGAGTTTGAGACCA
              GCCTGGCCAACATGGTGAAACCCCGTTTCTACTAAAAATACAAAAAGTTAGCTGGGCGTG
              [G,A]
              TGGCAGGCACCTGTAATCCCAGGTACTCGGGAGGCAGAAGTTGCAGTGAGCTGAGATTAT
              GCCACTGCACTCCAGCCTGGCTGACAGTGCGAAACTCCGTCTCAAAAAAAAAGAAAAAAG
              AAAAAAGAAATTCTAAATTCTGGGAGTTTTTCCATCAGTATCTGAGCAAGTTGGCAGGA
              AAGTTGAAAGAATGAAAGGAGACATGCCCAGGGCACCTGCTGGGAGAGTGAGTGGGGCTC
              AGGTAGCAGAGCCCTTTCCCAGGATGATAACCTCCTTGCCGTTGGTTGCAGAGAGGGGAC

30907         TGGCCCTGATGTTTTCTGGATTGGGATAGAATGGAAAGCTTCCTAAAAATGTTACTCTTT
              TCAACTCTTAGGATAGGGGTGCTGAAAGAAAAGGGAGAGACTATGGGTGGGTCCAATTCT
              TGTCTGTTTAAAAAGAAAATTCCGGCCGGGTGCAGTGGCTCATGCCTGTAATCTCAGCCT
              TTGGGAAGCCAAGGCGGTTGAATCACGAGGTTAGGAGTTTGAGACCAGCCTGGCCAACAT
              GGTGAAACCCCGTTTCTACTAAAAATACAAAAAGTTAGCTGGGCGTGGTGGCAGGCACCT
              [G,A]
              TAATCCCAGGTACTCGGGAGGCAGAAGTTGCAGTGAGCTGAGATTATGCCACTGCACTCC
              AGCCTGGCTGACAGTGCGAAACTCCGTCTCAAAAAAAAAGAAAAAAGAAAAAAGAAATT
              CTAAATTCTGGGAGTTTTTCCATCAGTATCTGAGCAAGTTGGCAGGAAAGTTGAAAGAAT
              GAAAGGAGACATGCCCAGGGCACCTGCTGGGAGAGTGAGTGGGGCTCAGGTAGCAGAGCC
              CTTTCCCAGGATGATAACCTCCTTGCCGTTGGTTGCAGAGAGGGGACACTGACCTCACCT

31447         CCTTTCCCAGGATGATAACCTCCTTGCCGTTGGTTGCAGAGAGGGGACACTGACCTCACC
              TTCTTCTCCGAGGACTGTTTTCACTTCTCAGACCGCGGGCATGCCGAGATGGCCATCGCA
              CTCTGGAACAACATGGTGAGCAGCCAAGGGCCTGGTGGGCCTTGTCAAGGGGGGGATCTAA
              GGATATTGACACTCTGTCTCACAATGGCAAAACTACTGGAGACATGGCTCCTTTCTCCCC
              AAAGCCCAAAGTGGCAGCACACCTTATTGGTCCTGATAGATTAATTCCAAAGGGAAAATA
              [C,A]
              CCTATATTTATCCAACACCCTTTGAAAGTTATACAAACACACACTCACACAACTTTATTC
              TTTGTTCCTTCAGCAATGCCCAGGTACTGCGAGGGGATCCCTTTGTAATCAGATAGGTTG
              GCTAGATGAAAATACCAACTTCTACCTCGTACTGTGTGACCTTGGGCAAACGATCTCTCT
              GGCCACCTGTATCAACATCTATAAAACAGTGAAAACAAGACAGGTCTCAGACAACGCATT
              GAGATCATGTGTACATGGCACCTAGCACAATAGTTAGCACTCAGCAAATGTCACCACCAT

31603         GGGCCTTGTCAAGGGGGGGATCTAAGGATATTGACACTCTGTCTCACAATGGCAAAACTAC
              TGGAGACATGGCTCCTTTCTCCCCAAAGCCCAAAGTGGCAGCACACCTTATTGGTCCTGA
              TAGATTAATTCCAAAGGGAAAATACCCTATATTTATCCAACACCCTTTGAAAGTTATACA
              AACACACACTCACACAACTTTATTCTTTGTTCCTTCAGCAATGCCCAGGTACTGCGAGGG
              GATCCCTTTGTAATCAGATAGGTTGGCTAGATGAAAATACCAACTTCTACCTCGTACTGT
              [G,A]
              TGACCTTGGGCAAACGATCTCTCTGGCCACCTGTATCAACATCTATAAAACAGTGAAAAC
              AAGACAGGTCTCAGACAACGCATTGAGATCATGTGTACATGGCACCTAGCACAATAGTTA
              GCACTCAGCAAATGTCACCACCATCAGCCTTCCAAGCACTCCGGGCTCAACTCATACCCA
              ACTCATTTCTCTAAACATCGAAAAGTGGAGATCCACACAGCCTGTTTTCCGAGGCTGATA
              CCTATTCCAGTCCTTTCTGATGGGAAGAAGGGACCTTATGAAATGAACATACAGTCTGGG

31685         CCAAAGCCCAAAGTGGCAGCACACCTTATTGGTCCTGATAGATTAATTCCAAAGGGAAAA
              TACCCTATATTTATCCAACACCCTTTGAAAGTTATACAAACACACACTCACACAACTTTA
              TTCTTTGTTCCTTCAGCAATGCCCAGGTACTGCGAGGGGATCCCTTTGTAATCAGATAGG
              TTGGCTAGATGAAAATACCAACTTCTACCTCGTACTGTGTGACCTTGGGCAAACGATCTC
              TCTGGCCACCTGTATCAACATCTATAAAACAGTGAAAACAAGACAGGTCTCAGACAACGC
              [A,G]
              TTGAGATCATGTGTACATGGCACCTAGCACAATAGTTAGCACTCAGCAAATGTCACCACC
              ATCAGCCTTCCAAGCACTCCGGGCTCAACTCATACCCAACTCATTTCTCTAAACATCGAA
              AAGTGGAGATCCACACAGCCTGTTTTCCGAGGCTGATACCTATTCCAGTCCTTTCTGATG
```

```
          GGAAGAAGGGACCTTATGAAATGAACATACAGTCTGGGGGTCTTTCAGGGACACCTGCCT
          GGTGCTTCCACTCTGCCTTCTGTGGCTGGCCACCAGCAACTGAACGGTTTCCGCACAGCA

31833     ACTGCGAGGGGATCCCTTTGTAATCAGATAGGTTGGCTAGATGAAAATACCAACTTCTAC
          CTCGTACTGTGTGACCTTGGGCAAACGATCTCTCTGGCCACCTGTATCAACATCTATAAA
          ACAGTGAAAACAAGACAGGTCTCAGACAACGCATTGAGATCATGTGTACATGGCACCTAG
          CACAATAGTTAGCACTCAGCAAATGTCACCACCATCAGCCTTCCAAGCACTCCGGGCTCA
          ACTCATACCCAACTCATTTCTCTAAACATCGAAAAGTGGAGATCCACACAGCCTGTTTTC
          [C,T]
          GAGGCTGATACCTATTCCAGTCCTTTCTGATGGGAAGAAGGGACCTTATGAAATGAACAT
          ACAGTCTGGGGGTCTTTCAGGGACACCTGCCTGGTGCTTCCACTCTGCCTTCTGTGGCTG
          GCCACCAGCAACTGAACGGTTTCCGCACAGCACTTGACCTGTCACCCCCAACAACTGGAT
          CCTCTTGCACGGAGCAAATGAAATGCCTTCCCAACCCAATGGTTTCTTTTAATCCAGGCT
          CAGTGGGTAACACAATCCCCACCCCAACCTGTATGTTCCCTCCTTTGTCCTATGACAACT

31970     GGTCTCAGACAACGCATTGAGATCATGTGTACATGGCACCTAGCACAATAGTTAGCACTC
          AGCAAATGTCACCACCATCAGCCTTCCAAGCACTCCGGGCTCAACTCATACCCAACTCAT
          TTCTCTAAACATCGAAAAGTGGAGATCCACACAGCCTGTTTTCCGAGGCTGATACCTATT
          CCAGTCCTTTCTGATGGGAAGAAGGGACCTTATGAAATGAACATACAGTCTGGGGGTCTT
          TCAGGGACACCTGCCTGGTGCTTCCACTCTGCCTTCTGTGGCTGGCCACCAGCAACTGAA
          [C,T]
          GGTTTCCGCACAGCACTTGACCTGTCACCCCCAACAACTGGATCCTCTTGCACGGAGCAA
          ATGAAATGCCTTCCCAACCCAATGGTTTCTTTTAATCCAGGCTCAGTGGGTAACACAATC
          CCCACCCCAACCTGTATGTTCCCTCCTTTGTCCTATGACAACTAAACAAGCTACATTCCA
          GCTCCTTTTATCACAGTTTCAGGCCCGTAGTGTCTCTGCCAACCACCGCTGTGCAAACGT
          TCCCACCCCTGTCAGCTCATCCAGTATGTCCAGCATCCCACTCGGCTGACTCACAATATT

32177     CCTTATGAAATGAACATACAGTCTGGGGGTCTTTCAGGGACACCTGCCTGGTGCTTCCAC
          TCTGCCTTCTGTGGCTGGCCACCAGCAACTGAACGGTTTCCGCACAGCACTTGACCTGTC
          ACCCCCAACAACTGGATCCTCTTGCACGGAGCAAATGAAATGCCTTCCCAACCCAATGGT
          TTCTTTTAATCCAGGCTCAGTGGGTAACACAATCCCCACCCCAACCTGTATGTTCCCTCC
          TTTGTCCTATGACAACTAAACAAGCTACATTCCAGCTCCTTTTATCACAGTTTCAGGCCC
          [A,G]
          TAGTGTCTCTGCCAACCACCGCTGTGCAAACGTTCCCACCCCTGTCAGCTCATCCAGTAT
          GTCCAGCATCCCACTCGGCTGACTCACAATATTGACTTTCTCCTTAGCTATACCATCTCC
          TCCTCTCTAGCAACCTCTTCTTTTAAGAACAGCATGTAAACTGGCTTTATCCTTGGCCTA
          GTTAATGGCAGACTCAGCTTATGTCGACTTCCATTGTCAGGGGGTTTTCCTCCTGTGGAC
          ATCACGTACCTGCCCACTCCAAGAACTTCTATTGTACTCTTTCAGCCCAAGACTCCGGAT

33018     AAGGTATGGCCTTCCTACCAGGTGGCACTCCAAGTCTGCTTAAATCTGGGACCCTCCAGG
          AATCTCCTGGGGCTGGATAGCCATAGTGACGGCTGGAACATGAAAAAGAGTCCATTGGTT
          TCTTTTCTTGTGAATTAACAATGTAGCTCTGGCCAGGCACGGTGGCTCATGCCTGTAATC
          CCAGCACTTTGGGAGGCCGAGGCAGGTGGATCGCTTGAGCCCAGGAATTAGACACCAACC
          TGGGCAACACAGGGGAGATTCTGTCTCTACAAAAATAATCAAAATATTAGCCAGGTGTGG
          [T,C]
          GGTGCATGCCTGTAGTCCCAGCTGCTCAGAAGGCTGACGTGAGAAGATCACTTGAGCATG
          GGAGGTCAAGGCTGCAATGAGCCGAGATGGCACCACCGCACTCCAGCCTGGGCAATAGAG
          TGAGACCCTATATCTCAAAAAACAAATAGAAAAAAAAAATATATGTAGCTCTGGCCTTCT
          CTTCTAAAGCAGTTCAGTAGCTCTTCCCATTCACCCAGGTAAGAGGCCTTTATTTCATAA
          AGATAAGTGGGAGGAGTTTAGATATGAAAACAAAACGTAAACACCGCACTGGAGCTATTG

33090     CTGGATAGCCATAGTGACGGCTGGAACATGAAAAAGAGTCCATTGGTTTCTTTTCTTGTG
          AATTAACAATGTAGCTCTGGCCAGGCACGGTGGCTCATGCCTGTAATCCCAGCACTTTGG
          GAGGCCGAGGCAGGTGGATCGCTTGAGCCCAGGAATTAGACACCAACCTGGGCAACACAG
          GGGAGATTCTGTCTCTACAAAAATAATCAAAATATTAGCCAGGTGTGGTGGTGCATGCCT
          GTAGTCCCAGCTGCTCAGAAGGCTGACGTGAGAAGATCACTTGAGCATGGGAGGTCAAGG
          [T,C]
          TGCAATGAGCCGAGATGGCACCACCGCACTCCAGCCTGGGCAATAGAGTGAGACCCTATA
          TCTCAAAAAACAAATAGAAAAAAAAAATATATGTAGCTCTGGCCTTCTCTTCTAAAGCAG
          TTCAGTAGCTCTTCCCATTCACCCAGGTAAGAGGCCTTTATTTCATAAAGATAAGTGGGA
          GGAGTTTAGATATGAAAACAAAACGTAAACACCGCACTGGAGCTATTGTGGAAACAAAAC
          AAGACTGTCCATGGTTCCCCAGCCATTATTATCTCAGCCATACCCCGAATTTCAAAATAA
```

FIGURE 3, page 26 of 33

33993   ATTACAGACCAGGTTTTTCTAGTCCTTTCCTGGTGACCTGGGCATGCCACCACCCTCCCC
ACTGCTCCCAACCTGATAAGCACATATATACCCGGTGAATTCATGTCTCACAATTAGAGT
CCTATGACATAGTGTCTGCAGGCTTTGGCTGATGTTCCCATAGTGTCTGCAGGCTTTGGC
TGATGTTCCCAGGGTTCCCTACTAGGAAGCAAAAAGCACCTTAAACTATTTCATCTTATT
TCATCTCCTGCCCCTCCTCTCACGTCCTTCTCGAGACTTTTGCAAAGGCAAAGCCAGAAG
[C,T]
TCCAGCAGCACCAGGGGATATTTTCCTCTTCCTCTGCCTTCTTCTGTCTTCTTATCTGAA
GAAGTTTCTCTTTCCCGAGGCCTAGTCCTCTACTGCTGCCTCTACTCCCTCTTCTGCAGA
AATCCTGCTCTCAGCCAGTGTTTGTATCTCCCCAGGTGCTGGGTGACAGCTCCAGCCTCC
TAACTGACATCCCTGTCTTCAGACTTAGAGCTCTTAGAATCGTGACTCTCAGCTCTGGCT
GCATATTAGAATCATTCAGGGACATTGTGTATGTGTGTATGTATGTGTATATATGTATGA

34284   AGCCAGAAGCTCCAGCAGCACCAGGGGATATTTTCCTCTTCCTCTGCCTTCTTCTGTCTT
CTTATCTGAAGAAGTTTCTCTTTCCCGAGGCCTAGTCCTCTACTGCTGCCTCTACTCCCT
CTTCTGCAGAAATCCTGCTCTCAGCCAGTGTTTGTATCTCCCCAGGTGCTGGGTGACAGC
TCCAGCCTCCTAACTGACATCCCTGTCTTCAGACTTAGAGCTCTTAGAATCGTGACTCTC
AGCTCTGGCTGCATATTAGAATCATTCAGGGACATTGTGTATGTGTGTATGTATGTGTAT
[G,A]
TATGTATGAATGTGTGTGTATGTGTGTGTGTGTATGTATGTATGTGTATGTGTGTATGTA
TGTATGTATGTATGACAGAGTCTCACTCTGTTGCCCAGGTTGGAGAGCAATGGCACCATC
TCAGTTCACTGCAACCTCCGTCTCCTGGATTCAAGCGATTCTCCTGCCTCAGTCTCCCAA
GTAGCTGGGGTTATAGGTGCATGCCACCATGACCAGCTAATTTTTGTATTTTTAGTAGAG
ACAAGGTTTCGCCATGTTGGCCAGGCCAGGCTGGTCTTTAACTCCTGACCTCAGGAGATC

34314   TTTTCCTCTTCCTCTGCCTTCTTCTGTCTTCTTATCTGAAGAAGTTTCTCTTTCCCGAGG
CCTAGTCCTCTACTGCTGCCTCTACTCCCTCTTCTGCAGAAATCCTGCTCTCAGCCAGTG
TTTGTATCTCCCCAGGTGCTGGGTGACAGCTCCAGCCTCCTAACTGACATCCCTGTCTTC
AGACTTAGAGCTCTTAGAATCGTGACTCTCAGCTCTGGCTGCATATTAGAATCATTCAGG
GACATTGTGTATGTGTGTATGTATGTGTATATATGTATGAATGTGTGTGTATGTGTGTGT
[G,A]
TGTATGTATGTATGTGTATGTGTGTATGTATGTATGTATGTGATGACAGAGTCTCACTCTG
TTGCCCAGGTTGGAGAGCAATGGCACCATCTCAGTTCACTGCAACCTCCGTCTCCTGGAT
TCAAGCGATTCTCCTGCCTCAGTCTCCCAAGTAGCTGGGGTTATAGGTGCATGCCACCAT
GACCAGCTAATTTTTGTATTTTTAGTAGAGACAAGGTTTCGCCATGTTGGCCAGGCCAGG
CTGGTCTTTAACTCCTGACCTCAGGAGATCCACCCACCTCGGCCTCCCAAAGTGCTGGGA

35392   GACGTCAGCATGATGGTGTAGAGACTCACTGGGGGATGAATAGTCCTGGAAGAAGGTGGA
AAGGGGCTTTTGAGGACTATAATAGTCTGTTGCCTGACTGGATGCTGGTATGTTCATTTT
ATCGAAACTTATCTGTTGCTCACTTATGATTTGTACTCGTTTCTATGTGTATGTTAGCTT
CAATTAAAAGTTTACTTGAGGCCGGGTACAGTGGCTCACACCTGTAATCCCAGCACTTTG
GGAGGCCGAGGCAGGCAGATCCCCTGAGGTCAGGAGTTCAATACCAGCCTAGCCAACATG
[A,G]
TGAAACCCCATCTCTACTAAAAATACAAAATTAGCCAAGCGTGGTGGCACGTGCCTATAA
TTCCAGCTACTTGGGAGGCTGAGACAGGAAAATCGCTTGAAACCAGGAGGCAGGGGTTGC
AGTGAGCCAAGATTGCATCATTGCACTCCAGCCTGGGTGACAAGAGTAAAACTCTGTCTC
AAATTTAAAAAAAAAAAAAAAAAAAAGTTTACTTGAAAAACAATATCAGTGCCTGACCGG
GCTTATCCCCAGAGAGTCTGACTTAATTGGTCTGGAGTGCGAGCTGGATTCGGTACTTTG

35599   ACAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCCCCTGA
GGTCAGGAGTTCAATACCAGCCTAGCCAACATGATGAAACCCCATCTCTACTAAAAATAC
AAAATTAGCCAAGCGTGGTGGCACGTGCCTATAATTCCAGCTACTTGGGAGGCTGAGACA
GGAAAATCGCTTGAAACCAGGAGGCAGGGGTTGCAGTGAGCCAAGATTGCATCATTGCAC
TCCAGCCTGGGTGACAAGAGTAAAACTCTGTCTCAAATTTAAAAAAAAAAAAAAAAAAAA
[T,A,G]
TTTACTTGAAAAACAATATCAGTGCCTGACCGGGCTTATCCCCAGAGAGTCTGACTTAAT
TGGTCTGGAGTGCGAGCTGGATTCGGTACTTTGTGAAAGCTCCTGAGATTATTTTAATGT
GCAGGGTTTATGAACCGCTGCCTTAGATCTGGTCCCCACAGAGAAATCAAGTAATCTGTA
TAAAAGAAAACCTGACCCAGTCACTCCCCTGCTTTCAAACTTCCAAAGCCTCCCACCTCT
GAAGGAGGCAGGCCAGGCCCCATAGCACAGCACACTAGGCCTCTGGGACTTGGCCTGGTT

35997   AGCTCCTGAGATTATTTTAATGTGCAGGGTTTATGAACCGCTGCCTTAGATCTGGTCCCC
ACAGAGAAATCAAGTAATCTGTATAAAAGAAAACCTGACCCAGTCACTCCCCTGCTTTCA
AACTTCCAAAGCCTCCCACCTCTGAAGGAGGCAGGCCAGGCCCCATAGCACAGCACACTA

```
        GGCCTCTGGGACTTGGCCTGGTTCACCTGATTAACCTCTCTGGCTACCATTTCCACCAGC
        GTCTGCCTCGCATGTTACAGTCTAGTGACTCCAGCAGCGTCCTGCACCACCTGTGGTGTT
        [C,A]
        CACACCTCTGCTAACTCTTGCTCTCCTCCTTCTCCTGGATTGCCCTTCTCACCTCCTTGC
        CCACTCCACCACTCAACTCAGGTGCCCACCTCCTGCAGGAAGCTACCTCTGAATCTCCAGG
        ACAGGCCAGTGGCCCACCCAGGTCCATTACACCCTGCCCAGTCCTGTCATTTGCTACGTG
        GTTGGTAGCCACAGTGCCTGGCTTAGGAAAGACTGGTTCTAGGAAAAACAATTTCATTCC
        CTGTGGCCAGCTCCAAGCCTTCCCCCGCCAAGCTTCTCCATTCAGGTCTCTGTGAATTTA

36085   GAAAACCTGACCCAGTCACTCCCCTGCTTTCAAACTTCCAAAGCCTCCCACCTCTGAAGG
        AGGCAGGCCAGGCCCCATAGCACAGCACACTAGGCCTCTGGGACTTGGCCTGGTTCACCT
        GATTAACCTCTCTGGCTACCATTTCCACCAGCGTCTGCCTCGCATGTTACAGTCTAGTGA
        CTCCAGCAGCGTCCTGCACCACCTGTGGTGTTCCACACCTCTGCTAACTCTTGCTCTCCT
        CCTTCTCCTGGATTGCCCTTCTCACCTCCTTGCCCACTCCACCACTCAACTCAGGTGCCA
        [C,T]
        CTCCTGCAGGAAGCTACCTCTGAATCTCCAGGACAGGCCAGTGGCCCACCCAGGTCCATT
        ACACCCTGCCCAGTCCTGTCATTTGCTACGTGGTTGGTAGCCACAGTGCCTGGCTTAGGA
        AAGACTGGTTCTAGGAAAAACAATTTCATTCCCTGTGGCCAGCTCCAAGCCTTCCCCCGC
        CAAGCTTCTCCATTCAGGTCTCTGTGAATTTAATTAATTCATCCATCCATCAAACAAGTA
        TTTACTGAGCACTAATATGTGCTAGGTACTGCTCCAGGTGCTGAGGACTCAGCAGTGAAA

36270   GCAGCGTCCTGCACCACCTGTGGTGTTCCACACCTCTGCTAACTCTTGCTCTCCTCCTTC
        TCCTGGATTGCCCTTCTCACCTCCTTGCCCACTCCACCACTCAACTCAGGTGCCACCTCC
        TGCAGGAAGCTACCTCTGAATCTCCAGGACAGGCCAGTGGCCCACCCAGGTCCATTACAC
        CCTGCCCAGTCCTGTCATTTGCTACGTGGTTGGTAGCCACAGTGCCTGGCTTAGGAAAGA
        CTGGTTCTAGGAAAAACAATTTCATTCCCTGTGGCCAGCTCCAAGCCTTCCCCCGCCAAG
        [C,T]
        TTCTCCATTCAGGTCTCTGTGAATTTAATTAATTCATCCATCCATCAAACAAGTATTTAC
        TGAGCACTAATATGTGCTAGGTACTGCTCCAGGTGCTGAGGACTCAGCAGTGAAAAGATG
        ACTGCTACTCTCATGGGACATACAGGATAGTAGGGAAAAGACAGATAATCAACAAGGTCA
        TTTCTGACCACATCTGTGGTTTAAGAAAAAGTCAAGCAGAGTGATGTGATACAGAGTAAT
        GGTGGGGGAGAGGGAGGCCTCCCTGAAGAAGTGACAGTGAATTGAGAAGCGCATGTCAAG

36481   GGTAGCCACAGTGCCTGGCTTAGGAAAGACTGGTTCTAGGAAAAACAATTTCATTCCCTG
        TGGCCAGCTCCAAGCCTTCCCCCGCCAAGCTTCTCCATTCAGGTCTCTGTGAATTTAATT
        AATTCATCCATCCATCAAACAAGTATTTACTGAGCACTAATATGTGCTAGGTACTGCTCC
        AGGTGCTGAGGACTCAGCAGTGAAAAGATGACTGCTACTCTCATGGGACATACAGGATAG
        TAGGGAAAAGACAGATAATCAACAAGGTCATTTCTGACCACATCTGTGGTTTAAGAAAAA
        [G,A]
        TCAAGCAGAGTGATGTGATACAGAGTAATGGTGGGGGAGAGGGAGGCCTCCCTGAAGAAG
        TGACAGTGAATTGAGAAGCGCATGTCAAGGGGGTTGCCAGGCAGAGGAAATAGGACCCACA
        TGGGCCTAGAGTCAGGAGTGAGCTTGAAGTGTCTGAGGAACTTAAAGGCCAATGTGACCA
        GAGGGAAGTGAACAAGGTGAAAAAGTTGGGCAGGGGCCAGGTCCCTAGATGCTTCTAAGC
        AGTAGAGTGATATGCTCTGGCTTACCCCTGGGTCCGTGTACCCTGGACTGGAAGAAAGCA

36619   ACAAGTATTTACTGAGCACTAATATGTGCTAGGTACTGCTCCAGGTGCTGAGGACTCAGC
        AGTGAAAAGATGACTGCTACTCTCATGGGACATACAGGATAGTAGGGAAAAGACAGATAA
        TCAACAAGGTCATTTCTGACCACATCTGTGGTTTAAGAAAAAGTCAAGCAGAGTGATGTG
        ATACAGAGTAATGGTGGGGGAGAGGGAGGCCTCCCTGAAGAAGTGACAGTGAATTGAGAA
        GCGCATGTCAAGGGGGTTGCCAGGCAGAGGAAATAGGACCCACATGGGCCTAGAGTCAGGA
        [G,A]
        TGAGCTTGAAGTGTCTGAGGAACTTAAAGGCCAATGTGACCAGAGGGAAGTGAACAAGGT
        GAAAAAGTTGGGCAGGGGCCAGGTCCCTAGATGCTTCTAAGCAGTAGAGTGATATGCTCT
        GGCTTACCCCTGGGTCCGTGTACCCTGGACTGGAAGAAAGCAAGGGTGGACCTGGAAAGA
        CCACTAGGAGGCTGCTGTTGATGGGTGAGAGAGGAAGGGGGCTGAGAGTAGGGTCAGGGC
        AGAGGAGGAGAGACGCTGTCGTGGGCTGGCGGATGGATGATGGGGAAGAGGAACAAAGGA

37088   GACCTGGAAAGACCACTAGGAGGCTGCTGTTGATGGGTGAGAGAGGAAGGGGGCTGAGAG
        TAGGGTCAGGGCAGAGGAGGAGAGACGCTGTCGTGGGCTGGCGGATGGATGATGGGGAAG
        AGGAACAAAGGATGACTTTTTGGTTTGGGGTCTAAGAAACTGGGTGGATGATTGAGCAGG
        TAGAGAAAAAATCAGCGTGGGAGGAAAAAAAATCAAGACTTCTGTTTTGGACATGGTGCA
        AACTGCCTTCCAGACATCCACATAGAGGTATCAGGATACAGAAGTTTGGAACTCACAGAG
        [G,C]
```

```
                 AAGTCAAGGCTGGAGATTGAAAAAAAAAAAAAAAAAAAAAAAAAGTGGGGTTATTAGCATA
                 GAGGGCCAATATGGTGAAACCCTGTCTCTACTGAAAATACAAAAATTATCCAGGCATGGT
                 GGCATGCACCTGTAATCCCAGCTACTCAGGGAGGCTGAGGCAGGAGAATTGCTTGAACCC
                 AGAGATGGGGTGGAGGTTGCAGTAAGCTGAGATCGTGCCATTGCACTCCAGCCTGGGTGA
                 CAGGGCAAGATTCCATCTAAAAAAAAAAAAAGCCACTACAGGATCAACTAAGAGCTCCTA

    37204        GAAGAGGAACAAAGGATGACTTTTTGGTTTGGGGTCTAAGAAACTGGGTGGATGATTGAG
                 CAGGTAGAGAAAAAATCAGCGTGGGAGGAAAAAAAATCAAGACTTCTGTTTTGGACATGG
                 TGCAAACTGCCTTCCAGACATCCACATAGAGGTATCAGGATACAGAAGTTTGGAACTCAC
                 AGAGGAAGTCAAGGCTGGAGATTGAAAAAAAAAAAAAAAAAAAAAAAAAGTGGGGTTATTA
                 GCATAGAGGGCCAATATGGTGAAACCCTGTCTCTACTGAAAATACAAAAATTATCCAGGC
                 [G,A]
                 TGGTGGCATGCACCTGTAATCCCAGCTACTCAGGGAGGCTGAGGCAGGAGAATTGCTTGA
                 ACCCAGAGATGGGGTGGAGGTTGCAGTAAGCTGAGATCGTGCCATTGCACTCCAGCCTGG
                 GTGACAGGGCAAGATTCCATCTAAAAAAAAAAAAAAGCCACTACAGGATCAACTAAGAGCT
                 CCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAGGCCAACTACCTAGCCCTGGGCAT
                 TCATTCCAGCTTTCAACTCCAGTGAGAGATGAGAAGGAGAGTGTGGAGGTAGATGGGAAA

    37485        ATACAAAAATTATCCAGGCATGGTGGCATGCACCTGTAATCCCAGCTACTCAGGGAGGCT
                 GAGGCAGGAGAATTGCTTGAACCCAGAGATGGGGTGGAGGTTGCAGTAAGCTGAGATCGT
                 GCCATTGCACTCCAGCCTGGGTGACAGGGCAAGATTCCATCTAAAAAAAAAAAAAAGCCAC
                 TACAGGATCAACTAAGAGCTCCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAGGCC
                 AACTACCTAGCCCTGGGCATTCATTCCAGCTTTCAACTCCAGTGAGAGATGAGAAGGAGA
                 [G,A]
                 TGTGGAGGTAGATGGGAAATGAGAAACAATGCTGTGTCCAGAGAGCTAAGAGAAGTCAGT
                 GTTTCAAGAGAGACAGAGCTGTCAACTTTGATGGATGCTTCTGAGAAGCCAAGCAAGTTG
                 AAGACAAAAAAAAAAAAAAATGATCTTTGGCTCTGCCCATATGGCGATCGTTGGTGGCCAG
                 GGCCAGAGCTTCCATCCAGCGATGGAGACTGCAGACTGGCTGGAGCGAGCAGCAGAGAGA
                 AGGAGAGATTAGGAAGTGCTGCCAGCACCTATAGACAGCTCTTCCCAGAAGTTATGAGAA

    37624        GGTGACAGGGCAAGATTCCATCTAAAAAAAAAAAAAAAGCCACTACAGGATCAACTAAGAGC
                 TCCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAGGCCAACTACCTAGCCCTGGGCA
                 TTCATTCCAGCTTTCAACTCCAGTGAGAGATGAGAAGGAGAGTGTGGAGGTAGATGGGAA
                 ATGAGAAACAATGCTGTGTCCAGAGAGCTAAGAGAAGTCAGTGTTTCAAGAGAGACAGAG
                 CTGTCAACTTTGATGGATGCTTCTGAGAAGCCAAGCAAGTTGAAGACAAAAAAAAAAAAAA
                 [-,A]
                 TGATCTTTGGCTCTGCCCATATGGCGATCGTTGGTGGCCAGGGCCAGAGCTTCCATCCAG
                 CGATGGAGACTGCAGACTGGCTGGAGCGAGCAGCAGAGAGAAGGAGAGATTAGGAAGTGC
                 TGCCAGCACCTATAGACAGCTCTTCCCAGAAGTTATGAGAAGTAACAGCCACGGTCACTG
                 GAGGGGACATGGATCAAAGAAAGGGCAGGTGAAGGAGGGGAGATGTCGGAGCAGGTTGTG
                 TACTGACGAGAAGGAACCAGTAGAAAGGGAGAAACTGATGCACTCATCAAACCCTTGTAA

    37685        CCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAGGCCAACTACCTAGCCCTGGGCAT
                 TCATTCCAGCTTTCAACTCCAGTGAGAGATGAGAAGGAGAGTGTGGAGGTAGATGGGAAA
                 TGAGAAACAATGCTGTGTCCAGAGAGCTAAGAGAAGTCAGTGTTTCAAGAGAGACAGAGC
                 TGTCAACTTTGATGGATGCTTCTGAGAAGCCAAGCAAGTTGAAGACAAAAAAAAAAAAAA
                 TGATCTTTGGCTCTGCCCATATGGCGATCGTTGGTGGCCAGGGCCAGAGCTTCCATCCAG
                 [C,T]
                 GATGGAGACTGCAGACTGGCTGGAGCGAGCAGCAGAGAGAAGGAGAGATTAGGAAGTGCT
                 GCCAGCACCTATAGACAGCTCTTCCCAGAAGTTATGAGAAGTAACAGCCACGGTCACTGG
                 AGGGGACATGGATCAAAGAAAGGGCAGGTGAAGGAGGGGAGATGTCGGAGCAGGTTGTGT
                 ACTGACGAGAAGGAACCAGTAGAAAGGGAGAAACTGATGCACTCATCAAACCCTTGTAAT
                 CACGATCATCTTCTGTGTGAATTAGTTCTGGGTTCCTGGAATAGCATCGGGAATCAGCCG

    37769        AGAGATGAGAAGGAGAGTGTGGAGGTAGATGGGAAATGAGAAACAATGCTGTGTCCAGAG
                 AGCTAAGAGAAGTCAGTGTTTCAAGAGAGACAGAGCTGTCAACTTTGATGGATGCTTCTG
                 AGAAGCCAAGCAAGTTGAAGACAAAAAAAAAAAAAAATGATCTTTGGCTCTGCCCATATGG
                 CGATCGTTGGTGGCCAGGGCCAGAGCTTCCATCCAGCGATGGAGACTGCAGACTGGCTGG
                 AGCGAGCAGCAGAGAGAAGGAGAGATTAGGAAGTGCTGCCAGCACCTATAGACAGCTCTT
                 [C,T]
                 CCAGAAGTTATGAGAAGTAACAGCCACGGTCACTGGAGGGGACATGGATCAAAGAAAGGG
                 CAGGTGAAGGAGGGGAGATGTCGGAGCAGGTTGTGTACTGACGAGAAGGAACCAGTAGAA
                 AGGGAGAAACTGATGCACTCATCAAACCCTTGTAATCACGATCATCTTCTGTGTGAATTA
```

FIGURE 3, page 29 of 33

```
          GTTCTGGGTTCCTGGAATAGCATCGGGAATCAGCCGCGCTGACCTTTAGCATTTATTCTG
          TCACTGTTACGATAGACTTGAGTTTCCTCAGTTCTTAAGAAAGTGGAAATAATACTACCT

38897     AGAGGGTCCTGGGATCAAAGGTATTTACACCCAGGGATATTTCAGATAAATCTTTTCATC
          TATGTGGAAAACATACAAAGTGGCGCAAGTGAGAAACTCCGATTTCCTAAGGTTGACAAG
          TCAAGTGCAGTAATGATGTCATGGTAACCAATATGTTTCCAAACTTTCCTAAGGTTGACT
          AGCCCCATGCACTTTGAGAAGTTGGTAAATAGGATTGTCGTCGTTTTATAAAATTGAAAA
          CACGGTGTCTTGCAATCACAGCCACTCACAAAGGAAGCCAGAGATGGTCCCAGCCCCTCC
          [G,A]
          CAGACTTCCTGTGGACTCAGGACTGGTGGTCTCTCCTGGGCCTTGCTGTACCCGGCAAAT
          CCAGGGGCACAGACTCAGGGTTCTGCCCTGCCGACAGATGCTGCCTAGCCTTCTGTGTGT
          CATAAGTCAACTCCCGCTCAGCCCCAGGCTGCTGGGTCCCTGCTGTGGGCCAAAAACCAG
          CCACTTCGCTGGTTTCTATCCCCCACCCCGTTCCCGAGGGAGGGGCTCTGGTGTGAGACA
          CCCCCTCAGAGAGGAAAGTGTCTCCCAGCTTTGGAGAGAATCGAGGTGTCCTTTCTCTCT

40155     AGACTGAGTGACAGAGTGAGACTGTCTCAAAAATAAAGTGCATCAAGCAGCTGTCCCGTG
          CCAGGCAGTATACTAGGATCTGGGGATCGGGGAGGCAAAGATAAAATAGACTCAGTGTCTG
          TTCCTGGAGCCTGCAATGGTCTTCCTCCCTCGCCACACCCACTGCCCTTGCCTGGCCCAC
          CTTCGAAGCCTGTGACTTGTCTCCCCAGCTCTCCTCTCCCTCTTCTCCATCCACCCTACA
          CTTGCTGCCAGACACAGATAGACCTTCCTGGAAATAACTTGCCCCATCAAGGCTGCTTGA
          [A,G]
          ATCCTTGCCTGATCCCTACTGCCCATTGACCAGAGTCTGGAGGGAGGGTCACCTCCCTCC
          ATGATACACACTGCACTCCTGGCCGGTGGATCCATCTCCCAGGAAGCCCCACGACTGCCC
          GCATCCAGGCCTTTCCTTTTGCCATCTGTTCCTGGAGGTTCATCTTCCATCTGCTATGAG
          AACATCCGCCTCCCTCCAGGTCCAGATGTTGCCTTTACTAAGCGATGGTTTCACCGTCTC
          TTACCTACCATTCCTGTCTCCAGACACTGACCCATGTGGGTCTCCTTTTCTATTTGTACC

40355     CTCCCCAGCTCTCCTCTCCCTCTTCTCCATCCACCCTACACTTGCTGCCAGACACAGATA
          GACCTTCCTGGAAATAACTTGCCCCATCAAGGCTGCTTGAAATCCTTGCCTGATCCCTAC
          TGCCCATTGACCAGAGTCTGGAGGGAGGGTCACCTCCCTCCATGATACACACTGCACTCC
          TGGCCGGTGGATCCATCTCCCAGGAAGCCCCACGACTGCCCGCATCCAGGCCTTTCCTTT
          TGCCATCTGTTCCTGGAGGTTCATCTTCCATCTGCTATGAGAACATCCGCCTCCCTCCAG
          [G,C]
          TCCAGATGTTGCCTTTACTAAGCGATGGTTTCACCGTCTCTTACCTACCATTCCTGTCTC
          CAGACACTGACCCATGTGGGTCTCCTTTTCTATTTGTACCTCTCATGAGACACCGACCCA
          GTCTCCTTTATGATGTGATTGTTTCTGCACATCTCAACTTCCTCCTGGGCCACAAGAAAA
          GATGTCACATCTTAACCCTCCAGTCTCATCACAGCTTCCAGCAAGGGGGCTAAACACAGC
          ACGTGCCCAATTCACATTCACTGAGAGGAGAGTGGAGAGGGGCATAGGAAGGCAAGAACG

40486     CAGAGTCTGGAGGGAGGGTCACCTCCCTCCATGATACACACTGCACTCCTGGCCGGTGGA
          TCCATCTCCCAGGAAGCCCCACGACTGCCCGCATCCAGGCCTTTCCTTTTGCCATCTGTT
          CCTGGAGGTTCATCTTCCATCTGCTATGAGAACATCCGCCTCCCTCCAGGTCCAGATGTT
          GCCTTTACTAAGCGATGGTTTCACCGTCTCTTACCTACCATTCCTGTCTCCAGACACTGA
          CCCATGTGGGTCTCCTTTTCTATTTGTACCTCTCATGAGACACCGACCCAGTCTCCTTTA
          [T,C]
          GATGTGATTGTTTCTGCACATCTCAACTTCCTCCTGGGCCACAAGAAAAGATGTCACATC
          TTAACCCTCCAGTCTCATCACAGCTTCCAGCAAGGGGGCTAAACACAGCACGTGCCCAAT
          TCACATTCACTGAGAGGAGAGTGGAGAGGGGCATAGGAAGGCAAGAACGCACACGATCTG
          CCCACATGCCTCCCCTCCCGGCCCTTCTGATTTGGGGATCTTTCATCTACTACAAAACCA
          GCTGTCCTTCCATGCTGCCCTTCCCTGATTTCTGGGTAGTCCTGGGATGGGAGAATGGGG

40512     CTCCATGATACACACTGCACTCCTGGCCGGTGGATCCATCTCCCAGGAAGCCCCACGACT
          GCCCGCATCCAGGCCTTTCCTTTTGCCATCTGTTCCTGGAGGTTCATCTTCCATCTGCTA
          TGAGAACATCCGCCTCCCTCCAGGTCCAGATGTTGCCTTTACTAAGCGATGGTTTCACCG
          TCTCTTACCTACCATTCCTGTCTCCAGACACTGACCCATGTGGGTCTCCTTTTCTATTTG
          TACCTCTCATGAGACACCGACCCAGTCTCCTTTATGATGTGATTGTTTCTGCACATCTCA
          [A,C]
          CTTCCTCCTGGGCCACAAGAAAAGATGTCACATCTTAACCCTCCAGTCTCATCACAGCTT
          CCAGCAAGGGGGCTAAACACAGCACGTGCCCAATTCACATTCACTGAGAGGAGAGTGGAG
          AGGGGCATAGGAAGGCAAGAACGCACACGATCTGCCCACATGCCTCCCCTCCCGGCCCTT
          CTGATTTGGGGATCTTTCATCTACTACAAAACCAGCTGTCCTTCCATGCTGCCCTTCCCT
          GATTTCTGGGTAGTCCTGGGATGGGAGAATGGGGACAGTTGTGACCACGAGGAAGCAGAG
```

40622   CCATCTGCTATGAGAACATCCGCCTCCCTCCAGGTCCAGATGTTGCCTTTACTAAGCGAT
GGTTTCACCGTCTCTTACCTACCATTCCTGTCTCCAGCACTGACCCATGTGGGTCTCCT
TTTCTATTTGTACCTCTCATGAGACACCGACCCAGTCTCCTTTATGATGTGATTGTTTCT
GCACATCTCAACTTCCTCCTGGGCCACAAGAAAAGATGTCACATCTTAACCCTCCAGTCT
CATCACAGCTTCCAGCAAGGGGGCTAAACACAGCACGTGCCCAATTCACATTCACTGAGA
[-,A,G]
GAGAGTGGAGAGGGGCATAGGAAGGCAAGAACGCACACGATCTGCCCACATGCCTCCCCT
CCCGGCCCTTCTGATTTGGGGATCTTTCATCTACTACAAAACCAGCTGTCCTTCCATGCT
GCCCTTCCCTGATTTCTGGGTAGTCCTGGGATGGGAGAATGGGGACAGTTGTGACCACGA
GGAAGCAGAGGTGGGAGTTCTACAGGCCCCACAGGGCTCTCTGCCATTGGTCACCTATCA
GTTCCCAATCTTTCAAAATCAGGTTTGATGGCCAAGGAAACGCTGGTGAGAAACCAAAAG

40654   GGTCCAGATGTTGCCTTTACTAAGCGATGGTTTCACCGTCTCTTACCTACCATTCCTGTC
TCCAGACACTGACCCATGTGGGTCTCCTTTTCTATTTGTACCTCTCATGAGACACCGACC
CAGTCTCCTTTATGATGTGATTGTTTCTGCACATCTCAACTTCCTCCTGGGCCACAAGAA
AAGATGTCACATCTTAACCCTCCAGTCTCATCACAGCTTCCAGCAAGGGGGCTAAACACA
GCACGTGCCCAATTCACATTCACTGAGAGGAGAGTGGAGAGGGGCATAGGAAGGCAAGAA
[T,C]
GCACACGATCTGCCCACATGCCTCCCCTCCCGGCCCTTCTGATTTGGGGATCTTTCATCT
ACTACAAAACCAGCTGTCCTTCCATGCTGCCCTTCCCTGATTTCTGGGTAGTCCTGGGAT
GGGAGAATGGGGACAGTTGTGACCACGAGGAAGCAGAGGTGGGAGTTCTACAGGCCCCAC
AGGGCTCTCTGCCATTGGTCACCTATCAGTTCCCAATCTTTCAAAATCAGGTTTGATGGC
CAAGGAAACGCTGGTGAGAAACCAAAAGAAGGTTCTAGCTGGGTGTTGACCTCTTTAGAG

40933   AGGGGCATAGGAAGGCAAGAACGCACACGATCTGCCCACATGCCTCCCCTCCCGGCCCTT
CTGATTTGGGGATCTTTCATCTACTACAAAACCAGCTGTCCTTCCATGCTGCCCTTCCCT
GATTTCTGGGTAGTCCTGGGATGGGAGAATGGGGACAGTTGTGACCACGAGGAAGCAGAG
GTGGGAGTTCTACAGGCCCCACAGGGCTCTCTGCCATTGGTCACCTATCAGTTCCCAATC
TTTCAAAATCAGGTTTGATGGCCAAGGAAACGCTGGTGAGAAACCAAAAGAAGGTTCTAG
[C,G]
TGGGTGTTGACCTCTTTAGAGGCCCATCCCGCTAAAGAGGGTTTGGGCACAGCCTAAATG
AGGGAGCTTTACAAAAGGGAAGCTCTGTGAAAACGTGCAGGGTTATCGCAGCATCTCAGG
AATGGGGACTAGGCAAGTCTTGGCTTGGTGATGGATGGTTCACGGAGATCCTTTCCACTG
ACCCCCGCTCCTCCTCCACAGGAGAGCCCTTACCTCTACACCCTGCGGAACAGCCGATTG
CTCCCAGACCAGGCTGAAGAAGCCCCCGAGGTGCTCTACTGGGCTGTCCCAGTGGCAGCG

41171   TCTTTCAAAATCAGGTTTGATGGCCAAGGAAACGCTGGTGAGAAACCAAAAGAAGGTTCT
AGCTGGGTGTTGACCTCTTTAGAGGCCCATCCCGCTAAAGAGGGTTTGGGCACAGCCTAA
ATGAGGGAGCTTTACAAAAGGGAAGCTCTGTGAAAACGTGCAGGGTTATCGCAGCATCTC
AGGAATGGGGACTAGGCAAGTCTTGGCTTGGTGATGGATGGTTCACGGAGATCCTTTCCA
CTGACCCCCGCTCCTCCTCCACAGGAGAGCCCTTACCTCTACACCCTGCGGAACAGCCGA
[T,C]
TGCTCCCAGACCAGGCTGAAGAAGCCCCCGAGGTGCTCTACTGGGCTGTCCCAGTGGCAG
CGGGAGTCGGCCTTGTGGTGGGCATCATCGGGACAGTGGTCTGGAGGTGCAGGAGAGGTG
GCCGGAGGGAAGATCCTCCAATGAGCCTGCGCACTGTGGCCCTCTAGGCCCGGGGGTGGG
TCCTCACCCTAAACTCCCTATAGCCACTCTCTTCACCGCCCTCTGCCCCAGCCACTCCCG
GCCACCAGGACATGCTTCAATGCCTGGTGCCATAGGAAGCCCAGGGACAGTCACAACTT

41379   TGGTGATGGATGGTTCACGGAGATCCTTTCCACTGACCCCCGCTCCTCCTCCACAGGAGA
GCCCTTACCTCTACACCCTGCGGAACAGCCGATTGCTCCCAGACCAGGCTGAAGAAGCCC
CCGAGGTGCTCTACTGGGCTGTCCCAGTGGCAGCGGGAGTCGGCCTTGTGGTGGGCATCA
TCGGGACAGTGGTCTGGAGGTGCAGGAGAGGTGGCCGGAGGGAAGATCCTCCAATGAGCC
TGCGCACTGTGGCCCTCTAGGCCCGGGGGTGGGTCCTCACCCTAAACTCCCTATAGCCAC
[T,C]
CTCTTCACCGCCCTCTGCCCCAGCCACTCCCGGCCACCAGGACATGCTTCAATGCCTGGT
GCCATAGGAAGCCCAGGGACAGTCACAACTTCTTGGGGCCTGGGCTTCTTCCAGGCCTA
TGCTCCTGGAATGGATACATTTAAATAAAGTCCAAAGCTATTTTATTCCTGGGTTTGCCT
GCGTGAAGCACTCACCTTCCATCTCTTGTGCAGCCCAGGTGTGGGAGCTGCCACTTTTTG
TGGCCTGCCTCCAGCAGGGCTGCCCAAGCCACGACCAACCAGAGCCCAAACTGCCTGCCA

41388   ATGGTTCACGGAGATCCTTTCCACTGACCCCCGCTCCTCCTCCACAGGAGAGCCCTTACC
TCTACACCCTGCGGAACAGCCGATTGCTCCCAGACCAGGCTGAAGAAGCCCCCGAGGTGC
TCTACTGGGCTGTCCCAGTGGCAGCGGGAGTCGGCCTTGTGGTGGGCATCATCGGGACAG

FIGURE 3, page 31 of 33

TGGTCTGGAGGTGCAGGAGAGGTGGCCGGAGGGAAGATCCTCCAATGAGCCTGCGCACTG
TGGCCCTCTAGGCCCGGGGGTGGGTCCTCACCCTAAACTCCCTATAGCCACTCTCTTCAC
[C,T]
GCCCTCTGCCCCAGCCACTCCCGGCCACCAGGACATGCTTCAATGCCTGGTGCCATAGGA
AGCCCAGGGGACAGTCACAACTTCTTGGGGCCTGGGCTTCTTCCAGGCCTATGCTCCTGG
AATGGATACATTTAAATAAAGTCCAAAGCTATTTTATTCCTGGGTTTGCCTGCGTGAAGC
ACTCACCTTCCATCTCTTGTGCAGCCCAGGTGTGGGAGCTGCCACTTTTTGTGGCCTGCC
TCCAGCAGGGCTGCCCAAGCCACGACCAACCAGAGCCCAAACTGCCTGCCACCACGAGCA

41880   ATCTCTTGTGCAGCCCAGGTGTGGGAGCTGCCACTTTTTGTGGCCTGCCTCCAGCAGGGC
TGCCCAAGCCACGACCAACCAGAGCCCAAACTGCCTGCCACCACGAGCATATCCTCAAGT
CACCAAACCCACTATTTCAAAGGCAGAAAAAATGCTGGTCACCAGGTGGTGGCTGGAATT
TTGGAGCTGGCTGGTTGCCATTCAGTCCAATCCAACACATACCTATTAAGCAACTGTTTT
GTATCCAGGACAATGCGAAGCACTGAGGTGCCTCCTAGGCTGTGCATGTCGCAGCCTGGC
[A,C]
GAGAGGTCAAACTCCTTCAATAACCAAGAAGCCACGTGATGATGTGTAACTACTAGGGCA
TCAGTAGGTAAATGTGTCTGATTGTTTTAAAGAATAGAAAGGGTTCTTCGGGGAAAGTTT
CTTGGGGGAGAGCAACCTTCACATGTCATTTTGGGAAAAGGAATAAAAAATGATTGGGAC
ACAAATACCTCCTATATTCTCAACCTGATTTTCTCAAGGTGCTAAATTTAGGAAAAAATT
CCTATTTCTATATGCCCAGGTTTCTGAGGGAAAACTAGAGAGAGTCTGAAAATATGGGCT

42278   AAAGGGTTCTTCGGGGAAAGTTTCTTGGGGGAGAGCAACCTTCACATGTCATTTTGGGAA
AAGGAATAAAAAATGATTGGGACACAAATACCTCCTATATTCTCAACCTGATTTTCTCAA
GGTGCTAAATTTAGGAAAAAATTCCTATTTCTATATGCCCAGGTTTCTGAGGGAAAACTA
GAGAGAGTCTGAAAATATGGGCTGCATTCACTGAGCCCCTGCTAGGGGCGAGGCCCCGTG
CTGGAGGCCTTCCACAGATGGTCTCTTTTATGCTGCACAAAAGCCCAGGGAGGGGGTAAA
[G,A]
GGAAAATCTTTGAAAATAGAAGTGATGCTTGCGCAACACCGTGAATGTACTAAACGCCGC
GAATTGTTCCATTTAAAATGATTAATTGTGTATCATGTGAATTTCACTTCAATAAAAAAG
AATCCAGGGAGGTAGACATCATCTGCATTGTAAACCTCTCTCTGATCCTGAAGTCCGGGA
TGATAAAGAGCCTGAGTCACAATCCCGGATGCAACACTGAAATGCTGTGCCCTGAAGCTG
CCTTCGCCAGCCTGAGCCCAGTGTCCCAGGCTCTGCATCTGTAAAAACTGGAGTAAGAGT

42339   AGGAATAAAAAATGATTGGGACACAAATACCTCCTATATTCTCAACCTGATTTTCTCAAG
GTGCTAAATTTAGGAAAAAATTCCTATTTCTATATGCCCAGGTTTCTGAGGGAAAACTAG
AGAGAGTCTGAAAATATGGGCTGCATTCACTGAGCCCCTGCTAGGGGCGAGGCCCCGTGC
TGGAGGCCTTCCACAGATGGTCTCTTTTATGCTGCACAAAAGCCCAGGGAGGGGGTAAAA
GGAAAATCTTTGAAAATAGAAGTGATGCTTGCGCAACACCGTGAATGTACTAAACGCCGC
[G,A]
AATTGTTCCATTTAAAATGATTAATTGTGTATCATGTGAATTTCACTTCAATAAAAAAGA
ATCCAGGGAGGTAGACATCATCTGCATTGTAAACCTCTCTCTGATCCTGAAGTCCGGGAT
GATAAAGAGCCTGAGTCACAATCCCGGATGCAACACTGAAATGCTGTGCCCTGAAGCTGC
CTTCGCCAGCCTGAGCCCAGTGTCCCAGGCTCTGCATCTGTAAAAACTGGAGTAAGAGTA
CACATTTTGCTTATCTCACGGCGCTGCTGAAAAATAAGGAACCGTGTGTGAACCTCTAAC

42612   CAACACCGTGAATGTACTAAACGCCGCGAATTGTTCCATTTAAAATGATTAATTGTGTAT
CATGTGAATTTCACTTCAATAAAAAAGAATCCAGGGAGGTAGACATCATCTGCATTGTAA
ACCTCTCTCTGATCCTGAAGTCCGGGATGATAAAGAGCCTGAGTCACAATCCCGGATGCA
ACACTGAAATGCTGTGCCCTGAAGCTGCCTTCGCCAGCCTGAGCCCAGTGTCCCAGGCTC
TGCATCTGTAAAAACTGGAGTAAGAGTACACATTTTGCTTATCTCACGGCGCTGCTGAAA
[A,G]
ATAAGGAACCGTGTGTGAACCTCTAACTCTAAAATGCTGCACAACTGAAAATGGCCTTTT
TCCTCGGTGAAGAGTTGGGATAAGGCCCAGACTGTTGGGGAAGATGTGAGACCCAGAGAT
GAGTTTGGGGAAATGGGGTAATAACATATGGGTGGAGAGTGCCCGCCTTCCTCTCAGGGA
GGTTCATCACCTTATCTCTTTCTGTCACAACAGAGAACCCGGAGGACCTATACCCAGTTC
CGTGTTCTTCTGGGCTTCAGTGTCTGTTTCTATACAATGGGAACAGCATGCATTCCCCTG

42817   TGCCTTCGCCAGCCTGAGCCCAGTGTCCCAGGCTCTGCATCTGTAAAAACTGGAGTAAGA
GTACACATTTTGCTTATCTCACGGCGCTGCTGAAAAATAAGGAACCGTGTGTGAACCTCT
AACTCTAAAATGCTGCACAACTGAAAATGGCCTTTTTCCTCGGTGAAGAGTTGGGATAAG
GCCCAGACTGTTGGGGAAGATGTGAGACCCAGAGATGAGTTTGGGGAAATGGGGTAATAA
CATATGGGTGGAGAGTGCCCGCCTTCCTCTCAGGGAGGTTCATCACCTTATCTCTTTCTG
[T,G]

```
CACAACAGAGAACCCGGAGGACCTATACCCAGTTCCGTGTTCTTCTGGGCTTCAGTGTCT
GTTTCTATACAATGGGAACAGCATGCATTCCCCTGCTTTTTCCTATAGACTGGAAAACGT
GGTGACCAAGTCACACATCCCAGCTTATGCTCCCGGCTTAAGACAGTGTAACGACAAAGG
TAACCCTTACACTCCTGGTTTGAGACAGTATAACGACAAAGGTAACATAGGAAGTCAAGG
AGTTCGCTTCACCGCCCCTCCCCCCACCCCACCCTTTTTTTTTCCTGCAAGTTTCTATTC
```

FIGURE 3, page 33 of 33